Mating incompatibility genes in the fungal pathogen *Colletotrichum lentis*

A Thesis Submitted to the College of
Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Plant Sciences
University of Saskatchewan
Saskatoon

By
Ramanpreet Kaur

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work, or in their absence, by the Head of the Department or the Dean of the College in which my thesis was done. It is understood that any copying or publication or use of the thesis, in whole or in part, for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or in part should be addressed to:

Head of the Department of Plant Sciences
University of Saskatchewan
51 Campus Drive
Saskatoon, SK S7N 5A8

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 – 110 Science Place
Saskatoon SK S7N 5C

ABSTRACT

Colletotrichum lentis is a serious foliar and stem pathogen causing anthracnose in lentil, which can lead to major losses in crop productivity. This fungal pathogen is asexual under field conditions but can undergo sexual reproduction in the laboratory, where two mating incompatibility groups (mIG-1 and mIG-2) were identified. Based on the absence of *MAT1-1* and the presence of *MAT1-2* in isolates of both mIGs, it was previously determined that *C. lentis* does not follow the classical mating system of heterothallic ascomycetes, which is governed through the mating type (*MAT*) locus. Bulk segregant analysis of an ascospore-derived population had revealed one major QTL, *qClMAT3*, on chromosome 3 of the *C. lentis* genome associated with mating incompatibility, but it is devoid of classical *MAT* genes. It was therefore hypothesized that genes other than the classical *MAT* genes may be regulating compatibility of two isolates as the first step in mating and the development of sexual reproductive structures. Using *Colletotrichum* gene-specific parameters, a total of 106 genes were predicted at QTL *qClMAT3* through the gene prediction tool FGENESH. Blasting (BLASTx) against the genera *Colletotrichum*, *Neurospora* and *Podospora* to identify homologous genes, and a search of the literature for evidence of an involvement in various pathways regulating mating identified 28 genes with a role in mating. As *C. lentis* appears to follow a heterothallic incompatibility system the status of these 28 genes in isolates of the two mIGs was analyzed in terms of presence or absence and sequence polymorphisms. Presence or absence of genes was not correlated with mIG, nor were there polymorphisms that were characteristic for specific mIGs. To study the temporal dynamics of candidate genes in vegetative mycelium and at several sexual stages, four developmental stages of perithecia were determined based on morphological characters such as the shape and color of perithecia as well as the development of asci and ascospores. Gene expression studies in individual mycelia, a vegetative co-culture of two compatible isolates and in two developmental stages of perithecia through RT-qPCR of 11 most promising candidate genes revealed that two genes, *Cl_Nop2* and *Cl_IDC1,* showed differential expression in the vegetative mycelia of CT-21 and CT-30. *Cl_IDC1* plays a role in mRNA trafficking during chemotropic interactions in model species *P. anserina* and *N. crassa*, and *Cl_Nop2* plays a role in cell viability. Four genes (*Cl_HP3, Cl_SART1, Cl_Hpt* and *Cl_KH)* associated with signaling pathways during chemotropic interactions showed upregulation in the co-culture compared to other stages, except for *Cl_HP3*, which was also upregulated in stage 3 perithecia. Expression of

Cl_BING4CT associated with vegetative incompatibility and *Cl_HP9* involved in cell cycle was downregulated in the co-cultures. The expression of *Cl_Cro1*, a gene expressed mainly during the dikaryotic stage and ascospore maturation in *P. anserine*, and mutation of which resulted in defects in septum formation in that species, was downregulated in stage 2 perithecia as compared to other stages. Expression of *Cl_BING4CT*, *Cl_HP9* and *Cl_IDC1*, specifically in the mycelial co-cultures, identified them as the most promising candidates for regulating mating incompatibility, which requires further validation through a gene silencing approach or the creation of mutants followed by mating experiments.

ACKNOWLEDGEMENTS

It is with a deep sense of gratitude that I wish to place on record my sincere thanks to my supervisor Dr. Sabine Banniza for her keen interest, constant encouragement, expert guidance, motivation and untiring efforts throughout my master's program. I would also like to thank my advisory committee members, Dr. Randy Kutcher, Dr. Aaron Beattie and Dr. Bill Biligetu as well as my external examiner Dr. Nicholas Larkan for their time, insights and recommendations.

I am also grateful to the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this project. Many thanks to the Pulse Crop pathology group for providing all the training and academic support during my research and also in completing my experiments. Special thanks to Cheryl Cho, Li Li, Kun and Kamal for helping me in harvesting perithecia. A sincere thanks to Ann Harley, faculty and students to make me feel like home away from my family. I really appreciate their friendliness and kindness that have moved the hearts of many.

I bow my head to the Almighty and my family members (My loving grandparents, my biggest support - my mother and my brother and his wife and especially to my father, who is no longer with us but continues to inspire me within, giving me strength to face every challenge in life with positivity). I could not have come this far without their generous support, positive attitude, invaluable blessings, selfless but untold sacrifices for my better and bright career. My special thanks are reserved for my friends Avneer, Ahneer, Ridhi, Supreet, Prabhjot and Navgeet for their motivation and encouraging me not to lose hope in hard times. I am also grateful to my roommates for being there in all my ups and downs.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Chapter 1

1.0 Introduction

Lentil (*Lens culinaris* Medik subsp. *culinaris*) is an important legume crop in Western Canada. It is a self-pollinating, dicotyledonous, diploid (2n=14) pulse crop grown to diversify and expand crop rotations and to improve the economic returns of producers. As a member of the legume family, Leguminosae, lentil can fix a significant portion of its nitrogen requirement from the air when inoculated with the appropriate rhizobial inoculant. Being a rich source of proteins, vitamins and minerals such as iron, zinc and micronutrients, lentil seeds are perceived as a healthy component of the human diet (Khazaei *et al*., 2017). Lentil can also help in combating micronutrient malnutrition due to the possibility of biofortification with selenium. The consumption of lentils can contribute to a reduction in the risk of diseases such as cancer, diabetes and cardiovascular diseases due to the presence of polyphenols and other bioactive compounds (Kumar and Baojun, 2017).

Lentil is affected by a wide range of fungal pathogens. These pathogens are a major impediment to crop productivity and are found in all lentil producing regions of the world. Fungal diseases such as anthracnose caused by *Colletotrichum lentis*, ascochyta blight caused by *Ascochyta lentis*, botrytis grey mould caused by *Botrytis fabae* and *B. cinerea* and stemphylium blight caused by *Stemphylium botryosum* are some of the major lentil diseases throughout the world (Taylor *et al.,* 2007). Among these, anthracnose and ascochyta blight are currently a major concern on the Canadian prairies due to the presence of favorable environmental conditions for disease development in these regions. These diseases mainly affect the leaves, stems, pods and stems of the lentil plants and thereby reduce the crop's productivity and marketability.

Accounting for up to 70% yield losses, anthracnose is the most serious foliar and stem disease of lentil in Saskatchewan (Morrall and Pedersen, 1991). The disease is caused by the asexual stage of *C. lentis*. The isolates from lentil can cause disease on faba bean and field pea under field conditions with varying levels of disease severity (Gossen *et al*., 2009). Two pathogenic races were described in the population of *C. lentis* (Buchwaldt *et al.,* 2004). The pathogen is anamorphic in nature but isolates can be mated successfully under laboratory conditions (Armstrong-Cho and Banniza, 2006). Although its sexual stage has not been reported in the field,

a comprehensive understanding of the mating system is important because of the potential risk of genetic recombination in the field that could generate new races of *C. lentis* and lead to resistance break-down in lentil cultivars.

Colletotrichum lentis appears to be heterothallic in nature based on the identification of two mating incompatibility groups (mIG-1 and mIG-2) in the *C. lentis* population as well as classical mating studies (Armstrong-Cho and Banniza, 2006, Menat *et al*., 2012). A complete understanding of the mating system of *C. lentis,* however, is still lacking as the mating system of *C. lentis* is more complex than the classical ascomycete mating system governed by the two mating type loci, *MAT1-1* and *MAT1-2*. In *C. lentis,* only *MAT1-2* is present in isolates of both mIGs, whereas *MAT1-1* appears to be completely absent (Menat *et al*., 2012). The genome of *C. lentis* is 56.10 Mb and consists of 10 core and 2 minichromosomes (Bhadauria *et al*., 2019). To reveal the genetic determinants of mating incompatibility, one single major QTL, *qClMAT3,* was identified based on composite interval mapping with a binary mating trait and SNP markers. However, *MAT1-2* is not present in this QTL indicating that genes other than classical *MAT* genes may be regulating mating incompatibility (Bhadauria *et al*., Dept. of Plant Sciences, University of Saskatchewan, unpublished).

1.1 Project hypothesis

Based on the presence of two mating incompatibility groups and the absence of *MAT1-2*, it was hypothesized that the genes underlying the QTL *qClMAT3* may be regulating sexual compatibility between mycelia of isolates of different mIGs.

1.2 Project Objectives:

The specific research objectives of this study were:

1. To identify all genes at QTL *qClMAT3*;
2. To identify genes regulating mating incompatibility based on published data on *Colletotrichum* spp. and other ascomycete model species;
    2.1 To identify and select candidate genes regulating incompatibility,
    2.2 To detect the presence or absence of candidate genes in isolates of mIG-1 and mIG-2,

2.3 To identify sequence polymorphisms in candidate genes in isolates of mIG-1 and mIG-2,

2.4 To identify the location of *MAT1-2-1* in the *C. lentis* genome;

3. To profile the expression of candidate genes in co-cultures and multiple perithecial developmental stages.

Chapter 2

2.0 Literature Review

2.1 Lentil Anthracnose

Anthracnose is a foliar and stem disease found in most lentil producing areas in western Canada. The disease was first reported in the Red River Valley of Manitoba in 1987 (Morrall, 1988). Over the next few years, the disease started spreading westwards and reached Saskatchewan in 1990. Since then, it has developed into the most serious foliar disease of Canadian-grown lentil.

Anthracnose develops under high relative humidity and at temperature higher than 15°C (Chongo and Bernier, 2000). Under favorable weather conditions and in the presence of inoculum, the symptoms of the disease become visible at the 10-12-node or early flowering stages of lentil plants. The disease initially manifests itself as tiny yellow lesions on the lower parts of stems and leaves. As lentil plants mature, these spots then enlarge into lesions and the disease starts spreading to the upper parts of plants. These lesions are tan to light brown in color, mostly dark at the borders and contain large numbers of small, irregularly shaped, brown to black structures known as microsclerotia, which give these lesions a black appearance. Finally, conidia develop in inconspicuous acervuli. Stem girdling and leaf senescence occur in case of the most adversely affected plants. Defoliation and stem girdling inhibit water and nutrient uptake and cause the plant to wilt resulting in areas of brown and dying plants in the field.

Microsclerotia on lentil debris in dust and soil particles serve as the primary source of inoculum for anthracnose development (Buchwaldt *et al*., 2018). Damage to microsclerotia caused by temperature fluctuations, repeated wetting and drying and soil microbial degradation is reduced by high melanin content in their cells, which enables the pathogen to remain viable for up to 3 years in the soil or infected plant debris buried in the soil. However, the pathogen causes less infection after 1 year when exposed to extreme weather conditions on the soil surface such as temperature extremes and ultraviolet radiation (Buchwaldt *et al.*, 1996). The spread of infective particles to neighboring fields mainly occurs through wind dispersal. Microsclerotia are mostly

carried along with dust to the neighbouring plants and fields during mechanical harvest. Secondary spread of the disease to the upper parts of the stem and leaves and neighbouring plants occurs by rain-splashed conidia formed in acervuli on stems and in leaf lesions (Buchwaldt, 2011). In contrast, seed transmission of this disease is mostly absent or occurs at a very low rate, and there is no evidence for the transmission of this disease from seed to seedling under field conditions (Morall, 1997). Taken together, it appears that infected lentil debris and dust play a major role in the spread of disease.

Yield losses caused by anthracnose has led to the implementation of various management measures for disease prevention. These measures range from chemical (e.g. chlorothalonil) and cultural (e.g. crop rotations) methods to the employment of resistant cultivars. In order to reduce the primary source of inoculum in fields, lentil growers are advised to follow 3- to 4-year crop rotations and practise reduced tillage. In order to protect the crop from secondary spread of the disease, fungicide treatments are applied after identification of early symptoms such as premature leaf drop caused by the disease. Fungicide applications with picoxystrobin, mancozeb, fluxapyroxad, isofetamid, chlorothalonil or azoxystrobin as major active ingredients has been recommended at the 10-12-node stage to early flowering with a second application after 10-14 days (Guide to Crop Protection, 2020). In addition, farmers can also make use of a fungicide decision support system (FDSS) to evaluate the severity and risk of the disease. FDSS also guides the farmers in identifying the appropriate timing for fungicide applications on the basis of four risk factors, (A) plant density, (B) number of days with rain in the last 14 days, (C) rain in the 5-day weather forecast, and (D) early symptoms of anthracnose (Buchwaldt *et al*., 2018).

In view of the potential yield losses caused by anthracnose and the high cost of fungicides, there has been an ongoing effort to identify sources of anthracnose resistance in lentil. While working on cultivated lentil accessions obtained from various gene banks, Buchwaldt *et al*. (2004) identified the two pathogenic races Ct0 and Ct1 of the causal organism *C. lentis* in the Canadian population based on differences in disease severity of lentil accessions. In order to avoid confusion with isolate identification numbers as well as designations of resistance and avirulence genes, these races were later re-designated as race 0 and race 1 (Armstrong- Cho *et al*., 2012). After screening 1771 lentil accessions, Buchwaldt *et al*. (2004) identified 16 accessions including PI 0345629, PI 320937 and Indianhead with resistance to race 1. These three accessions were subsequently exploited as sources of resistance to study the genetics, biology

and epidemiology of *C. lentis* (Buchwaldt *et al*., 2018). Accession PI 320952 (Indianhead) was used for the development of the first partially resistant cultivar CDC Robin (Vandenberg *et al.*, 2002). Many other varieties have been developed since then. Classical genetic studies revealed that resistance to race 1 in cultivated lentil is due to the combination of recessive and dominant genes (Buchwaldt *et al.*, 2013). To date, three dominant genes, *CtR3*, *CtR4* and *CtR5*, and two recessive genes, *ctr1* and *ctr2*, have been identified in progenies developed by crossing resistant lines with the susceptible cultivar Eston (Tullu *et al*., 2003, Buchwaldt *et al.*, 2013).

Landrace VIR421 (Buchwaldt and Diederichsen, 2004), as well as accessions of the wild lentil species *L. ervoides* and *L. lamottei* (Tullu *et al.*, 2006) were identified as sources of partial resistance to race 0. Gela *et al.* (2020) showed that the eastern and central European accessions of *L. culinaris* previously shown to have resistance to race 0 (Shaikh *et al.*, 2013) only showed low levels of partial resistance and have higher susceptibility to race 0 isolates when compared with the interspecific line LR-59-81 derived from *L. ervoides*, thus limiting the sources of resistance to race 0 to wild lentil species, particularly *L. ervoides*. Various biotechnological as well as advanced breeding methods have been employed to transfer resistance from wild lentil species into the cultivated lentil and to reveal the genetics of resistance. In a study undertaken by Tullu *et al.* (2013), resistance genes from *Lens ervoides* accession IG 72815 were transferred into cultivated lentil through embryo rescue and 29% of the lines in the resulting F_8 population showed resistance to race 0. Two recessive genes were implicated in conferring resistance; however, segregation distortion in interspecific populations as a result of variable levels of fertility in subsequent populations obstructed the mapping of quantitative trait loci (QTLs) associated with resistance to race 0 and race 1. An intraspecific recombinant inbred line (RIL) population, LR-66, was developed from *L. ervoides* accessions L01- 827A and IG 72815 and was phenotyped for resistance to race 0 and race 1 in F_9 RILs (Bhadauria *et al*., 2017). Together with genotypic single nucleotide polymorphism (SNP) data, an analysis of QTLs for anthracnose resistance was conducted and revealed five loci showing significant association with resistance to race 0 and six with resistance to race 1. Currently, identification of candidate genes for resistance is underway with the objective to develop specific markers to select individuals derived from interspecific crosses with resistance to both races.

2.2 *Colletotrichum*

2.2.1 The Genus *Colletotrichum*

Owing to highly uncertain and confusing taxonomy, the ascomycete genus *Colletotrichum* has been subject to several revisions over the years. With earlier studies depicting *Colletotrichum* to be closely related to *Verticillium* (a member of the Plectosphaerellaceae), more comprehensive morphological and phylogenetic studies revealed that *Colletotrichum* is a member of the Glomerellaceae of the order Glomerellales in the subclass Hypocreomycetidae of the Sordariomycete (Reblova *et al*., 2011). Most species of the genus were classified into complexes on the basis of the genealogical concordance phylogenetic species recognition (GCPSR) concept followed by phylogenetic analysis. A total of 11 species complexes, namely *caudatum*, *graminicola*, *spathianum*, *destructivum*, *acutatum*, *dematium*, *gigasporum*, *gloeosporoides*, *bininense*, *truncatum* and *orbiculare* (Marin-Felix *et al*., 2017) with three more recent complexes (*dracaenophilum*, *magnum* and *orchidearum*) (Damm *et al*., 2019) have been described to date. This classification was based on multilocus phylogenetic analyses carried out with the help of gene sequences for actin (act), chitin synthetase (chs1), glyceraldehyde 3 phosphate dehydrogenase (GAPDH) and DNA lyase (apn2) and morphological characterization in terms of conidial shape and size, the shape of appressoria, and the presence or absence of setae.

The genus *Colletotrichum* includes species with different lifestyles, which are primarily known as phytopathogenic species that infect various crops worldwide causing major losses in crops. *Colletotrichum* spp. can infect a wide range of legumes, cereals, fruits and vegetable crops. In addition to anthracnose of field crops, *Colletotrichum* spp. can also cause many other diseases in crops such as sugarcane and coffee. Species in the genus cause diseases at both preharvest (on aerial parts of plants such as stems, leaves, seedlings and tubers) and postharvest stages (in fruits and vegetables). *Colletotrichum* spp. such as *C. lindemuthianum*, *C. gloeosporioides*, *C. graminicola* and *C. higginsianum* have been widely used as models in studying fungal infection structures, their morphogenesis, fungal-plant interactions, pathogenicity and host defence mechanisms (O'Connell and Panstruga, 2006). Due to its pervasiveness, mode of destruction and scientific importance as a model for the study of host-pathogen interactions, *Colletotrichum* is ranked among the top 10 genera of scientific and economic importance (Dean *et al*., 2012).

Apart from being phytopathogenic, some *Colletotrichum* spp. are also human pathogens; for example, *Colletotrichum gigasporum* has been reported to cause human infections in the form of subcutaneous cyst (Liu *et al*., 2014). Some of the *Colletotrichum* spp. also show saprophytic (Huang *et al*., 2013) as well as endophytic (Hiruma *et al.,* 2016) lifestyles. *Colletotrichum tofieldiae* is an endophytic species that improves phosphorus uptake of *Arabidopsis thaliana* and promotes better growth under phosphorus-deficient conditions (Hiruma *et al.,* 2016). Other examples of endophytic species in the genus include *C. karstii, C. colombiense*, *C. parsonsiae* and *C. boninense*.

Species of this genus display different levels of host specificity and can either infect a single host or can infect multiple hosts. The three species *C. orchidearum*, *C. plurivorum* and *C. sojae* from the *C. orchidearum* species complex have a wide range of hosts. For example, *C. sojae* can infect species of the genera *Glycine*, *Medicago*, *Phaseolus and Vigna* (Fabaceae,) but also *Amaranthus* (Amaranthaceae) and *Capsicum* (Solanaceae) (Damm *et al.,* 2019). The isolates of *C. lentis* have narrow host range preferences, restricted mainly to the tribe Fabeae of the Fabaceae. Host range studies undertaken by Gossen *et al.* (2009) related to *C. lentis* under controlled and field conditions confirmed that isolates of this pathogen can infect lentil, field pea and faba bean resulting in varying levels of disease severity. There was no significant difference in patterns of virulence between the two pathogenic races 0 and 1 of *C. lentis*. Symptoms of the disease were also observed in chickpea varieties when inoculated with *C. lentis* under controlled conditions; however, no symptoms of the disease were reported on chickpea in field conditions. *Colletotrichum lentis* was also found to cause anthracnose in common vetch plants (Xu *et al*., 2017).

2.2.2 *Colletotrichum lentis*

Originally, the species causing anthracnose on lentil was determined to be *C. truncatum* (Morrall, 1988), but subsequent comparative studies of the morphological and multilocus molecular data of lentil isolates indicated that the isolates belong to the distinct, new species *C. lentis* of the destructivum clade (Damm *et al*., 2014).

Two pathogenic races described in the western Canadian population of *C. lentis* by Buchwaldt *et al*. (2004) were associated with genetic variation on the basis of polymorphism in the repeated lengths of minisatellites in intergenic spacer regions (IGS) of the ribosomal DNA

(Durkin *et al.*, 2015). Sequence alignment of 39 nucleotides of each minisatellite showed a varying number of repeats in race 0 and race 1 *C. lentis* isolates. Amplification of 39 nucleotides of minisatellites in 50 *C. lentis* isolates from lentil with known race identity using a PCR-based probe confirmed that race 0 isolates have 2 or 4 repeats and race 1 isolates have 7 or 9 repeats, thus repeat numbers can be used to differentiate isolates of the two races. It was speculated that this polymorphism could be the result of unequal crossing over events during mitosis and meiosis, or mutation (Durkin *et al.*, 2015). A set of 15 candidate effectors were identified from 2000 expressed sequence tags (ESTs) generated from *C. lentis*-infected lentil leaf tissues. A single nucleotide polymorphism was identified in the two candidate effectors *ClCE6* and *ClCE8* on the basis of which Kompetitive Allele Specific PCR marker (KASPar) assays were developed to differentiate isolates of race 0 from race 1. The *ClCE6* KASPar marker showed polymorphism between isolates of the two pathogenic races when tested on 52 randomly selected *C. lentis* isolates with known race identity, suggesting its use in determining race identity of *C. lentis* isolates (Bhadauria *et al*., 2015). Furthermore, the host-specific toxin *ClToxB* was identified as a virulence factor, which probably contributes to virulence differences between race 0 and race 1 isolates of *C. lentis*. This was demonstrated through *in planta* expression analysis where the expression level of *ClToxB* peaked at 48 hours per inoculation (biotrophy-necrotrophy switch) and transcript levels of the toxin were higher in lentil plants infected with the isolate CT-30 of the more virulent race 0 than those infected with isolate CT-21 of the less virulent race 1. However, no sequence polymorphism was detected between the races (Bhadauria *et al*., 2015).

Menat *et al*. (2012) identified two mating incompatibility groups, mIG-1 and mIG-2, in the *C. lentis* population with complete lack of selfing. Molecular population studies revealed that mIG-1 was always associated with race 0 but not with race 1 in all field isolates indicating the possibility of association of mating incompatibility group with virulence (Menat *et al*. 2016). However, both traits occurred in equal proportions and all possible combinations in an ascospore-derived *C. lentis* population developed under laboratory conditions, confirming that both traits were inherited independently and the association was an effect of linkage disequilibrium due to clonal reproduction. Random mating was also rejected based on gametic disequilibrium indices calculated from amplified fragment length polymorphism (AFLP) markers of field isolates.

Colletotrichum lentis is hemibiotropic in nature. The pathogen undergoes a biotrophy to necrotrophy switch to progress from a symptomless biotrophic phase to a destructive necrotrophic phase. Studies undertaken by Armstrong-Cho *et al*. (2012) revealed that there is no difference in the infection process of the two races as isolates of both races progress from the biotropic to the necrotropic phase. However, the rate of conidial germination and appressoria formation is significantly higher for race 0 isolates than race 1 isolates.

Colletotrichum lentis infects aerial parts of the plant. After spore deposition on lentil plants and conidial germination, which takes place between 3 to 6 hours post inoculation (hpi), appressoria are differentiated. The formation of appressoria is the prerequisite for host invasion and takes places between 6-12 hpi (Chongo *et al.,* 2002). Penetration pegs develop on the underside of appressoria and penetrate into the epidermal cell wall of the host. The penetration leads to the formation of infection vesicles underneath the penetration sites that enlarge into thick primary hyphae. The primary hyphae remain confined to first infected epidermal cells (Armstrong-Cho *et al*., 2012). This is the biotrophic phase. Then the primary hyphum differentiates thin secondary hyphae at 48-68 hpi, which invade neighboring cells and the fungus enters the necrotrophic phase. During this phase, cell death occurs and manifests with visible symptoms in the form of lesions on stems and leaves, which will eventually lead to defoliation and wilting of the plant. Conidia found in acervuli in anthracnose lesions restart the disease cycle (O'Connell *et al*., 2000).

2.3 Sexual reproduction in Ascomycetes

2.3.1. Importance of genetic recombination

Fungal species show wide genetic variability and rapid evolution. This can be the result of various mechanisms such as transposable elements or mutations, and recombination events due to sexual reproduction or parasexuality. Sexual reproduction is an important phenomenon in almost all organisms and creates genetic diversity by exchange of genetic information, thus contributing to specis evolution. Sexual reproduction in fungi is a very complex cellular process that takes place under specific environmental and genetic conditions necessary for the expression of the genes regulating the process. Genetic recombination created through sexual reproduction in a pathogen can lead to resistance breakdown in its host, thus confering advantages to the pathogen for its survival even in adverse environments through elimination of harmful genes and

by combining beneficial alleles. However, in case of clonal reproduction, the advantageous mutations located at independent loci are unable to recombine and the pathogen may be unable to evolve with changes in environments (Moller and Stukenbrock, 2017). In some of the fungal species, pathogenicity and mating are linked and mating is considered an important step in the lifecycle. For instance, in the smut fungus *Ustilago maydis*, the dikaryotic structures required for the growth of the fungus are the result of mating and if mating does not take place, the organism remains avirulent (Feldbrugge *et al*., 2004). At the genetic level, there are various genes, proteins and metabolic pathways that regulate the sexual process in fungal species. The genes harboured at the mating type (*MAT*) locus are the major regulators of sex in ascomycete fungi and also determinants of the mating strategies and are thus part of the mating or breeding systems (Debuchy *et al*., 2010). In general, ascomycetes have two types of mating system: homothallism (self-fertile) and heterothallism (self-sterile) (Glass *et al*., 1988).

2.3.2. Sexual cycle in ascomycetes

Sexual reproduction in ascomycetes is diverse and involves different mating strategies in different species. Sexual reproduction occurs inside the multicellular fruiting bodies emerging from mycelia (Taylor, 1995). Differentiation of the female reproductive structures (ascogonia) and male gametes (spermatia or antheridia) under appropriate environmental conditions initiates the sexual cycle. The female gametangium (ascogonium) is a multinucleate cell with an apical receptor element, the trichogyne, which grows towards the male cell. The male cell can be a specialised hypha (antheridium) or asexual vegetative spore (microconidum, macroconidium or spermatium). During the fertilization process, the male nucleus enters the ascogonium with the help of the trichogyne. In heterothallic species, mating only occurs in the presence of a partner of the opposite mating type. However, in homothallic species, mating occurs in a single isolate as the mating type genes are present in the same nucleus. Apart from this, the subsequent steps of sexual reproduction are identical in both mating systems. The fertilization of male and female gametes initiates the development of the fruiting body. The fruiting bodies are known by different names, such as cleistothecium, apothecium, perithecium or pseudothecium, depending on their morphology.

After plasmogamy, an ascogeneous hypha develops from the ascogonium. There is no fusion between the nuclei after fertilization; instead each nucleus proliferates and hook-shaped cells

develop, which are referred to as croziers. The parental nuclei move into the ascogenous hyphae and undergo mitosis in the crozier. After septa formation, the crozier gives rise to three types of cells: an upper binuclear cell and a lateral and basal uninuclear cell. The binuclear cell undergoes karyogamy. Immediately after the fusion of the nuclei of the dikaryotic cell, meiosis of the single diploid nucleus takes place and these meiotically derived haploid nuclei are separated into ascospores (Debuchy *et al*., 2010). However, in some of fungal species such as *Colletotrichum* spp., meiosis is followed by additional mitosis, thus eight haploid nuclei are formed after the mitosis of the four meiotically derived haploid nuclei.

2.3.3. Genetic basis of the mating process

Sexual compatibility in ascomycete fungi is mainly controlled by the genes encoded at the *MAT* locus. This locus harbours alternate alleles known as idiomorphs, which are differentiated on the basis of unique and unrelated sequences. Idiomorphs are commonly referred to as *MAT1-1* and *MAT1-2* (Turgeon and Yoder, 2000) and possess conserved sequences of unique identity. Idiomorphs consist of a minimum of one mating type gene, which encode transcription factors and also determine sexual compatibility. The *MAT1-1* idiomorph may have more than one gene but always contains the primary *MAT1-1-1* encoding for a protein that contains the conserved alpha domain of a DNA binding protein (Dyer *et al*., 2016). Similarly, *MAT1-2* possesses a corresponding gene *MAT1-2-1* encoding for a protein with conserved high mobility group (HMG). These conserved sequences have been exploited in various molecular studies to locate their positions in the genome of fungal isolates and to understand the mating systems of various species (Arie *et al*., 2000). The *MAT* genes are regulated by different transcription factors and control the sexual cycle. For instance, the gene *pro1* encodes a *MAT* regulating protein, which positively regulates the HMG box protein, and that, in turn, controls the expression of both *MAT1-1* and *MAT1-2* in *Podospora anserina* (Gautier *et al*., 2018). Likewise, the light-dependent sexual reproduction of *Trichoderma reesei* is regulated by the three regulatory protein-encoding genes *blr1*, *blr2* and *elv1* (Seibel *et al*., 2012). These genes regulate the expression of *MAT1-2-1* and hence play a role in the sexual cycle.

Apart from primary *MAT1-1-1* and *MAT1-2-1*, a number of secondary *MAT1-1* and *MAT1-2* genes have also been identified in various fungal species and were shown to play an important role in sexual reproduction. Their actual function in the sexual cycle may vary from species to

species. Structural analysis of the *MAT* locus in *P. anserina*, *N. crassa* and *Sachharomyces cerevisiae* revealed functional roles of the mating genes during the fertilization event. In *P. anserina*, *MAT1-1-1* and *MAT1-2-1* have been identified as the genes responsible for mating specificity. Mutation experiments within the *MAT* locus revealed two other genes, *MAT1-1-2* and *MAT1-1-3*, in the *MAT1-1* idiomorph (Coppin *et al*., 1993). It was found that *MAT1-1-1* has its role during fertilization, whereas *MAT1-1-2* and *MAT1-1-3*, along with *MAT1-1-1* are also required for nuclear identity, cellular division, and development and maturation of the sexual structures. Recently, expression analysis of the mating type genes in *Botrytis cinerea* revealed that the secondary *MAT* genes *MAT1-2-4* and *MAT1-1-5* are also essential for sexual reproduction (Rodenburg *et al*., 2018). *MAT1-2-4* of *B*. *cinerea* is involved in ascomatal development whereas *MAT1-1-5* is involved in protoascomatal development. Similarly, *MAT1-2-4* plays a role in fruiting body development of *Sclerotinia sclerotiorum* (Doughan and Rollins, 2016) and it is essential for heterokaryon formation in *Aspergillus fumigatus* (Yu *et al*., 2017).

Chemical signals in the form of pheromones were reported to be involved in the mating process and are responsible for recognition and attraction of partners of opposite mating types, as well as the fusion of female gametes with male cells. The pheromone-encoding genes have been described in various ascomycete fungi such as *N. crassa*, *S. cerevisiae*, *P. anserina* and various *Fusarium* species. The two pheromones MFa and MFα factor pheromones were the first to be identified in the two mating types of *S. cerevisiae*, which then became a model for studies related to pheromones (Nakayama *et al*., 1985). The pheromone-encoding genes are regulated by the *MAT* genes and are expressed in a mating type-dependent manner. Expression studies in various fungal species such as *N. crassa* and *P. anserina* have demonstrated that *MAT1-1* regulate the MFα type pheromone genes whereas *MAT1-2* individuals express the MFa type pheromone genes whereas in case of homothallic species, both are expressed in the same individual. As revealed by gene expression studies in *N. crassa*, these pheromones are typically expressed in either spermatia or conidia and are essential for male fertility in heterothallic species (Kim and Borkovich, 2006). These a- and α-type pheromones are recognised by receptors encoded by *pre-1* and *pre-2* genes, respectively. These G coupled receptors are present in the female structures and initiate a mitogen activated protein (MAP) kinase signal transduction pathway, thus activating other network pathways and leading to sexual reproduction. These receptors control expression of pheromones and female fertility in *N. crassa*. In addition to their primary role in

mate recognition, pheromones also play a role in post-fertilization events. For instance, pheromones and their receptors are required for the production of ascospores in *N. crassa.* Expression of the α-factor was downregulated in *MAT1-1* by deletion of the a-factor receptor and deletion of the α-factor receptors resulted in the a-factor downregulation in *MAT1-2* (Kim *et al*., 2012).

Genes encoding pheromones and their receptors have also been found in the homothallic species *Sordaria macrospora* and *Fusarium graminearum.* Interestingly, a low level of self-fertility was still achieved in the homothallic species *Sordaria macrospora* after knocking out both pheromone genes, but allowing expression of the two receptor genes, which suggests that pheromone receptor interactions may not be necessary to initiate signaling in this mating pathway (Mayrhofer *et al*., 2006). The presence of both pheromones and receptors may not be required for self-fertility but their presence may be required at later stages of development. Mutants with single deletion of α-factor receptors produces fewer ascomata than wild type isolates of *F. graminearum.* Additionally, outcrossing events were also promoted by deletion of the α-factor indicating its role in selfing in *F. graminearum.* However, the deletion of receptors along with the α-factor leads to a decrease in these outcrossing events (Lee *et al*., 2008). The defining features of the α-factor pheromone is its short length and presence of a conserved C-terminal CaaX domain with either a CVIL or CVVM sequence. The a-factor pheromones are recognized by the presence of a hydrophobic region and single peptide at the N-terminal.

Most of the heterothallic ascomycetes have a bipolar mating system in which mating is controlled by a single mating locus/two alternative alleles. Homothallic individuals contain alternate alleles in the same nucleus, do not require any mating partner and thus undergo selfing. Recombination and mutations can also lead to switching between homothallism and heterothallism. Studies undertaken by Yun *et al*. (1999) on different heterothallic and homothallic species in the genus *Cochliobolus* suggested that heterothallism was ancestral and that homologous recombination and unequal crossover events within the very short identical sequences of the idiomorphs led to homothallism. The evidence strongly supporting the evolution of homothallism from heterothallism is based on the structural organization and phylogenetic analysis of the *MAT* loci in *Cochliobolus* spp. It was predicted that the sequences of the *MAT* genes are highly conserved in heterothallic species whereas homothallic species have a unique arrangement of *MAT* genes suggesting the evolution from one type to the other (Yun *et*

al., 1999). Phylogenetic studies related to mating type loci suggested that the ascomycetous species *N. crassa* and *P. anserina* may have been derived from a common asexual ancestor, and that heterothallism was acquired concomitantly by both species (Coppin *et al*., 1997). Recently, a study conducted by Wilken *et al*. (2018) on mating type idiomorphs of heterothallic species of *Thielaviopsis* indicated an unexpected presence of the *MAT1-1-2* of the *MAT1-1* idiomorph at the *MAT1-2* locus. This could be the result of an aberrant crossover event which might have moved the *MAT1-1-2* gene from the *MAT1-1* to the *MAT1-2* idiomorph. The idiomorph organization seen in *Thielaviopsis* spp., highlighted the complex evolution of the sexual reproduction and the mating type locus in this genus. Similar results have also been obtained for the mating system of the heterothallic *Leptographium* spp. where the *MAT1-2* idiomorph possesses truncated *MAT1-1-1* genes homologous to those from the *MAT1-1* idiomorph, but lacking the conserved alpha domain (Duong *et al*., 2013). These studies provide evidence for the evolution of mating systems and indicate the high plasticity at the mating type loci, including the possibility for recombination in these areas.

Although mating type genes are highly stable in fungi in general, mating type switching has been reported in many ascomycetes such as *S. cerevisiae, S. pombe, Kluveromyces lactis* and *N. crassa.* In *S. cerevisiae*, mechanisms have evolved that led to mating type switching in the cells (Haber, 2012). Usually, mating type switching takes place due to the presence of silent copies of mating type genes in a species, and these become active during recombination events in the form of gene conversion. With the help of various recombination repair factors such as endonucleases and transposases the switching of mating types can occur. Various genes and transcriptional factors play an important role in the switching mechanisms in yeast and other ascomycetes. Mating type switching mechanisms vary in different species due to their different lifestyles. For example, in *S. cerevisiae*, *HO,* an endonuclease gene, is crucial for *MAT* switching, whereas in *K. lactis*, the specialized protein α3 is required for successful switching (Barsoum *et al*., 2010). Taken together, recombination in the form of crossovers or gene conversions within and around the *MAT* loci play an important role in the evolution of mating systems and hence in mating system determination.

A recent study on the mating system of diverse members of *Xylariales* indicated the absence of mating type genes in all members of the genus (Robinson and Natvig, 2019). *MATA_HMG* encoding genes detected in these species were shown to be orthologs of non-mating type genes

instead of mating type genes on the basis of BLAST searches of the predicted *Xylariales* protein sequences against the protein sequences of *MAT1-2-1* and *MAT1-1-3* from various *Neurospora* spp. Repeated attempts to find the homologs of the mating type genes in the members of *Xylariales* spp. including the use of *sla2* and *apn2* known to flank the mating regions of most species of the Ascomycota failed. Based on these results, it has been postulated that either the mating regions have been lost or they have been altered in *Xylariales,* and that sexual reproduction in these fungi is regulated in an atypical manner rather than the classical mating system of the ascomycetes studied to date.

2.3.4. Mating system and mating type genes in the genus *Colletotrichum*

Colletotrichum species are mostly anamorphic in nature (Sutton, 1992) and only a few species have been associated with the teleomorph state, previously referred to as *Glomerella*. The mating type locus in the genus is known as the *MAT* locus. Both, homothallic and heterothallic mating systems have been found among *Colletrotrichum* spp. The heterothallic *Colletotrichum* spp. do not follow the typical bipolar mating system according to which both mating type idiomorphs, *MAT1-1* and *MAT1-2*, need to be present for mating to take place. In many *Colletotrichum* spp., including *C. lindemuthanium (*García-Serrano *et al*., 2008), *C. graminicola* (Vaillancourt *et al.,* 2000) and *C. lentis* (Menat *et al.*, 2012), the MAT1-2 idiomorph has been identified with the help of the highly conserved HMG domain present in the idiomorph. However, all efforts to locate the *MAT1-1* in *Colletotrichum* spp., even in homothallic strains have failed*,* indicating that the *MAT1-1* locus probably plays no role in sexual incompatibility in *Colletotrichum* spp. Unusual mating systems have been reported for various *Colletotrichum* spp., such as *C. gloeosporioides* (syn. *G. cingulata*) and *C. graminicola* (syn. *G. graminicola*) (Cisar and Tebeest *et al*., 1999; Vaillancourt *et al*., 2000). This atypical mating system was believed to be the result of mating type switching with “unbalanced heterothallism” to be the primary cause (Wheeler, 1956).

All isolates of *C. lindemuthanium* had *MAT-1-2-1* in spite of being heterothallic (Rodriguez-Guerra *et al*., 2005). Hence, it has been suggested that mating in the genus is regulated in an atypical manner distinct from the classical system of ascomycetes. Genetic analysis of the *MAT1-2* idiomorph in various species indicated that *MAT1-2* genes are highly conserved in various isolates of *Colletotrichum* species (Crouch *et al*., 2014). In *C. gramincola*, the *MAT1-2-1*

gene showed 99.5% nucleotide similarity among different isolates with only four bases showing polymorphism. These polymorphic bases are located outside of the HMG domain. Of these bases, three are part of the first intron and polymorphism in the fourth base leads to a change from asparagine to aspartic acid. However, there is no polymorphism in the HMG domain of *MAT1-2-1*. Various molecular techniques have been employed to detect the presence of the *MAT1-1* idiomorph in the homothallic *C. graminicola* genome, but all attempts failed (Du *et al*., 2005).

2.3.5. Mating incompatibility and the mating system in *C. lentis*

Colletotrichum lentis appears to be asexual in nature as the sexual stage has only been observed under laboratory conditions (Armstrong-Cho and Banniza, 2006). The mating system of *C. lentis* is more complex than the classical mating system of ascomycetes (Menat *et al*., 2012). Two mating incompatibility groups, mIG1 and mIG2, have been described in *C. lentis* with no evidence of selfing. Molecular population studies undertaken by Menat *et al.* (2016) suggested that sexual reproduction occurs very rarely or is absent under field conditions in isolates of *C. lentis*.

Characterization of the *MAT* locus of *C. lentis* revealed the presence of only the *MAT1-2* idiomorph in isolates of the two mating incompatibility groups (Menat *et al.*, 2012), whereas the classical bipolar mating system of ascomycetes requires one partner to be *MAT1-1* and the other to be *MAT1-2*. The absence of *MAT1-1* was consistent with other *Colletotrichum* spp. (Vaillancourt *et al*., 2000). Sequence identity and the constitutive expression of *MAT1-2-1* in sexually compatible isolates CT-21 (race 1), CT-30 (race 0) and the co-culture of CT-21 and CT-30 indicated that the mating process in *C. lentis* is regulated by genes other than the *MAT* genes (Warale, 2015).

QTL mapping of an ascospore-derived population of *C. lentis* from a cross between field isolates CT-21 and CT-30 was conducted to reveal the genetic determinants of mating incompatibility. It was found to be associated with a single major effect QTL, *qClMAT3* with LOD score of 21.92 and R^2 value is 87.12, on chromosome 3 (Bhadauria *et al*., Dept. of Plant Sciences, University of Saskatchewan, unpublished). The reference genome was that of *C .lentis* isolate CT-30 belonging to mIG-1 (Bhadauria *et al*., 2019). The absence of *MAT1-2-1* in the QTL *qClMAT3*

indicated that genes other than the classical *MAT* genes may be regulating compatibility between mycelia of isolates of two mIGs.

Chapter 3

3.0 Material and methods

3.1 Identification of genes at QTL *qClMAT3*

Bulk segregant analysis of the ascospore-derived *C. lentis* population from a cross of CT-30 × CT-21 revealed the major QTL *qClMAT3* was associated with mating incompatibility (Bhadauria *et al.,* Dept. of Plant Sciences, University of Saskatchewan, unpublished). This QTL, mapped on linkage group 3 (LG3) of the *C. lentis* genome, was found by composite interval mapping performed with 931 SNP markers and the binary mating trait. This locus is approximately 350 kbp in length. In the current project, the sequence of the QTL *qClMAT3* interval was dissected for the prediction of genes underlying this QTL. Using *Colletotrichum*-specific gene-finding parameters, gene sequences were predicted *in silico* with the gene prediction tool FGENESH (Salamov and Solovyev, 2000). The predicted gene sequences were then blasted (BLASTx) against the NCBI non-redundant protein sequence database of the genus *Colletotrichum* and the fungal model species in *Neurospora* and *Podospora* to identify homologous genes in these species. All gene sequences with hits in these species were catalogued. Subsequently, the genome of *C. lentis* was accessed on KnowPulse (http://knowpulse2.usask.ca/portal/node/1) and the predicted gene sequences of QTL *qClMAT3* were blasted against the genome in order to identify the gene copy number.

3.2 Identification of candidate genes for mating incompatibility

Once the identification of genes underlying *qClMAT3* was completed, these genes were further researched to identify those with a role in mating incompatibility. The functional role of each domain-containing protein encoded by these genes was determined with the help of the NCBI reference sequences (Refseq), which provided information about the homologous domains and their functions through protein family and domain databases such as UniProtKB and InterPro.

The function of various domains was also determined by searching through the literature for evidence of their involvement in various pathways regulating mating.

3.2.1 Isolates

Twenty *C. lentis* isolates derived from the previously developed ascospore-derived mating population from a cross between field isolates CT-30 (mIG-1) × CT-21 (mIG-2) with known mating incompatibility group were available for use in the study (Table 3.1). Of these 20 isolates, half belong to mIG-1 and the other half belong to mIG-2. Presence or absence, and sequence identity of candidate genes among isolates of the two mating incompatibility groups were assessed for these 20 isolates. Two scenarios could indicate involvement in mating incompatibility: Candidate genes could be present in one mIG and absent in the other, or, if present in both, candidate genes could have identical sequences in isolates of the same mIG, but show sequence polymorphism when compared to isolates of the opposite mIG.

Table 3.1. *Colletotrichum lentis* isolates selected to assess the status of candidate genes involved in mating. Isolates were selected from an ascospore-derived mating population from a cross between field isolates CT-30 (mIG-1) × CT-21 (mIG-2) with known mating incompatibility groups (mIG).

mIG-1 isolates	mIG-2 isolates
GT-2	GT-7
GT-12	GT-52
GT-29	GT-69
GT-30	GT- 75
GT-64	GT-96
GT-99	GT-111
GT-105	GT-125
GT-119	GT-135
GT-120	GT-138
GT-164	GT-146

3.2.2 Presence or absence of candidate genes in isolates of two mIGs

Isolates were cultured on oatmeal-agar medium (OMA: 30 g oatmeal flour [Quick Oats Robin Hood, Smucker Food of Canada, Markham, Ontario, Canada], 8.8 g granulated agar [Molecular genetics, Fisher bioreagents™, Thermo Fisher Scientific, Ottawa, ON], 1 L distilled water) and incubated at 22°C for 7-10 days. To obtain mycelia for extraction of DNA, a small amount of fungal mycelium of each isolate was transferred from the OMA plates into flasks of 50 mL yeast glucose medium (1 g $NH_4H_2PO_4$, 0.2 g KCl, 0.2 g $MgSO_4.7H_2O$, 10 g glucose, 5 g yeast extract, 0.01 g $ZnSO_4.7H_2O$, 0.005 g $CuSO_4.5H_2O$, 1L distilled water). The liquid cultures were incubated on a shaker for 2-3 days at 21-23°C and 130 rpm. Mycelia of isolates were harvested by filtration using mira cloth, collected in 1.5 mL Eppendorf tubes, stored for 2 days at -80°C, and lyophilized for 1 day (Labconco Free Zone 7753022, Labconco Corp, Kansas City, MO). After adding one sterile solid glass bead (Fisher Scientific, Ottawa, ON) to each 2 mL microcentrifuge tube, the lyophilized mycelium of each isolate was pulverized with a Gene Grinder (SPEX SamplePrep 2010, Horiba) for 1 min at full speed. DNA was extracted using DNA extraction buffer (250 mL 5M NaCl, 200 mL 1M Tris -Cl, 50 mL 10% SDS, 50 mL 0.5 M EDTA, 450 mL distilled water). Subsequently, 500 µL of extraction buffer and 4 µL of RNase (100 mg mL^{-1}) were added to 100 mg ground samples and incubated at 65°C in a water bath for 20 min. The samples were centrifuged at 13,000 rpm for 10 min. The supernatant was collected in a 2.0 mL microcentrifuge tube that was incubated at -20°C for 10 min after treatment with an equal volume of ice-cold isopropanol. A second centrifugation was then performed at 13,000 rpm for 2 min and the supernatant was discarded without causing disturbance to the pellet. The pellet was washed with 70% ethanol and was centrifuged at 13,000 rpm for 2 min. The DNA pellets were air dried by discarding ethanol and keeping the tubes upside down on Kim wipes for an hour. The DNA pellets were dissolved in 40 µL sterile water and stored at -20°C. The quality and quantity of DNA samples was determined using the A260: A280 ratio (~ 1.8) on a Nanodrop ND8000 spectrophotometer (Thermo Fisher Scientific, Ottawa, ON).

The concentration of the DNA samples was diluted to 50 ng μL^{-1} with sterile distilled water for PCR. The primers of 28 candidate genes potentially involved in the pathway associated with mating incompatibility were designed from the sequence information obtained as described in Section 3.1 using the Primer3Plus tool Version 2.4.2 (Untergasser *et al*., 2012) (Table 3.2). The PCR amplifications were done using a Bio-Rad thermal cycler (C1000 Touch™ thermal cycler,

Bio-Rad Laboratories, Inc., Canada) with the final reaction volume of 20 µL. The reaction mixture for each gene consisted of 2 µL of 50 ng genomic DNA, 2 µL of 10x PCR buffer (Invitrogen, Carlsbad, CA), 0.6 µL of 50 mM $MgCl_2$, 0.5 µL of 20 µM dNTP (Invitrogen, Carlsbad, CA), 0.5 µL each of 10 mM forward and reverse primers, and 0.2 µL of 1 U Taq DNA Polymerase (Invitrogen, Carlsbad, CA). PCR reaction conditions consisted of an initial denaturation step at 94°C for 5 min followed by 40 cycles at 94°C for 45 s, then the annealing temperature of each primer (Tm) for 30 s and 72°C for 90 s, and a final elongation step at 72°C for 10 min.

The PCR products were stained with 6X DNA gel loading dye (Thermo Scientific, Ottawa, ON) and separated on a 1.5% agarose gel in 1x Tris-acetate-EDTA (TAE) buffer (pH 8.0) stained with Gel Red (Biotium, Inc, CA). The PCR products were visualized under UV light to detect the presence or absence of identified genes using Bio-Rad ChemiDoc™ Gel Imaging System (Bio-Rad Laboratories, Inc.). The size of the PCR products was estimated using the 1Kb Plus DNA ladder (Invitrogen, Carlsbad, CA).

Table 3.2. Forward and reverse primer pairs designed for determining the presence or absence, sequence polymorphism and expression analysis of selected candidate genes involved in mating incompatibility in isolates of *Colletotrichum lentis*.

Primer ID	Gene ID	Forward sequence	Reverse sequence	Amplicon size
Cl_5-qClMAT3	*Cl_HP9*	TAGACAGTTCCTGGCGACCT	GATGTCGAGGATCGAGGTGT	164bp
Cl_6-qClMAT3	*Cl_Nop2*	CAGCGTCAAGACGAACAAGA	TCTCGACAAGCCTGACATTG	209bp
Cl_7-qClMAT3	*Cl_Byr4*	GATATCGAGGACTGGGACGA	CACATGCGTCTCCTCGTCTA	168bp
Cl_8-qCLMAT3	*Cl_HP1*	AGAGCAACCCAGCTTGAAGA	GGATGCAGATGCCAAAAGTT	168bp
Cl_14-qClMAT3	*Cl_RNS3*	CAAGAGGACGATGTGAGCAA	ACTTGACTGCCCCTATGTGG	223bp
Cl_21-qClMAT3	*Cl_IDC1*	CATGCCTCTTCATCTCGTCA	CCACGAAAGTGTTTCCGATT	222bp
Cl_23-qClMAT3	*Cl_PAN1*	TACCAACGAGGGTTGTAGGC	CAGGAAGAGCAGGAATCGAC	197bp
Cl_24-qClMAT3	*Cl_HP7*	ACACCCACAAACGCAACATA	AGCCCCAAAGATGGACTTTT	173bp
Cl_33-qClMAT3	*Cl_PfPl*	CCATCGTCCCAGATATCACC	CTTCGTAGCTCATCGGGTTC	222bp
Cl_35-qClMAT3	*Cl_IMP*	TGGTTCTACCACCCTTCTCG	CCATTCTATCGGCCACTTGT	186bp
Cl_37-qClMAT3	*Cl_BHLH*	TCAAGCTGCCAGCACATTAC	GCACGTTTGAAATCCTCCAT	159bp
Cl_50-qClMAT3	*Cl_HECT*	ACGCAGGAGGTCTCAAGAAA	ATGGCCAAGCCAAGTACAAC	220bp
Cl_52-qClMAT3	*Cl_HP8*	CGGCAAAGATGGAGAGAAAG	GATGGTAGATGGGCTTTCCA	233bp
Cl_53-qClMAT3	*Cl_IPP1*	GTACAAGAACGCTCGCAACA	CTCGTTATCCACCACGTCCT	159bp
Cl_56-qClMAT3	*Cl_BING4CT*	ACTGTCGATCGAGGAGGAGA	CCTCGACCTCGTTGTTCTTC	249bp
Cl_59-qClMAT3	*Cl_SART1*	ACCGCACCTGTCTTATCACC	AATTCGATATGCAGGGCTTG	189bp

Primer ID	Gene ID	Forward sequence	Reverse sequence	Amplicon size
Cl_63-qClMAT3	*Cl_KH*	CGGCGACTACATCAAAGACA	TCTCTCGGCCATATTCATCC	248bp
Cl_64-qClMAT3	*Cl_Duf974*	TCAACCTCCCATTGAGCTTC	GAGCTCGAGTTTCTGGATCG	236bp
Cl_65-qClMAT3	*Cl_Denn*	CAACCCCTTGAGCGATATGT	GGACTCGGACATGGCATACT	222bp
Cl_73-qClMAT3	*Cl_IK13*	CGCTACGTCGTCTGTCGTTA	GTGAGTTGCTGGGGTCAAAT	218bp
Cl_75-qClMAT3	*Cl_Hpt*	TCGAGAAGTTGTCGCAGTTG	CTTCTTCTTCCGTGCCAAAG	243bp
Cl_80-qClMAT3	*Cl_MAK*	CGATCGGTCGACAACCTTAT	ACGTTGCTGACAGATTGACG	200bp
Cl_81-qClMAT3	*Cl_HP3*	ATAGCATAGCCGAAGCCTCA	CGTAAGGAACAATGGCCTGT	201bp
Cl_83-qClMAT3	*Cl_HP2*	GTCGGCTTCTCACCTCTGTC	CAAGTCATTGAGTCGCTGGA	177bp
Cl_91-qClMAT3	*Cl_RNS3*	AAGTTTCTCAGTGGCGAGGA	GAGACAGCCTCTCCATACGC	151bp
Cl_96- qClMAT3	*Cl_WD40*	GCTGGGAAGATGCTTGTAGC	TTGCCAGCAGAAACTTGTTG	151bp
Cl_102-qClMAT3	*Cl_Cro1*	GTACCGAGTGGCTCGAAGAG	GTGATTGCGGGTGAAATCTT	223bp
Cl_106-qClMAT3	*Cl_SNARE*	TTACGTGCTGCTTTCTGGTG	GAGGAGATCCGGAACATTGA	191bp
Cl_37_qpcr-qClMA	*Cl_BHLH*	AATCTGCCCTTCGGCTATG	TGGCATGTTGTAGTGATCTGC	153bp
Cl_5_qpcr-qClMAT	*Cl_HP9*	AACACCTCGATCCTCGACAT	AGAACTCCAGCGAAGACTGG	151bp
Cl_7_qpcr-qClMAT	*Cl_Byr4*	TGCTGGATCGAAGAGAGTCA	AGGGGTGATGAGTGAAATGG	170bp
Cl_75_qpcr-qClMA	*Cl_Hpt*	GTCGATGGACTCGGCCTTAG	GTGATGCGTTTGAGGCACAG	195bp

3.2.3 Sequence polymorphisms in candidate genes in isolates of mIG-1 and mIG-2

Sequence polymorphisms in genes present in all isolates irrespective of their mating incompatibility group were analyzed in the Genomics Workbench (QIAGEN CLC Genomics Workbench 20.0 (https://digitalinsights.qiagen.com)). All PCR products of the genes showing presence in all isolates were gel extracted from all 20 isolates using the Monarch DNA gel extraction kit (New England Biolabs, UK) following the manufacturer's protocol. According to this protocol, the gel band for the PCR product of each isolate was eluted and placed in a pre-weighed 1.5 mL microfuge tube. Gel-dissolving buffer equal to four volumes of the weight of the gel band was added and the tube was incubated in a water bath at 55°C for 15 min. After that, the sample was transferred to a spin column and centrifuged at 13,000 rpm for 1 min. The flow-through was discarded and the spin column was treated with 400 µL of Monarch DNA wash buffer containing 95% ethanol and centrifuged at 13,000 rpm for 1 min. The washing step was repeated twice. After that, 20 µL of the DNA elution buffer was added to the spin column containing the filtrate, centrifuged for 2 min at 13,000 rpm and the flow-through was transferred to 1.5 mL microfuge tubes and stored at -20°C until sequencing at Eurofins Genomics LLC, Kentucky.

Sequencing was performed bidirectionally using gene-specific forward and reverse primers and from the sequencing data, and consensus for each sequence was derived by aligning the sequences from both forward and reverse primers. DNA sequencing data were then compared *in silico* by aligning the sequences of the 20 isolates using the multiple sequencing alignment tool in the CLC Genomics Workbench 20.0 (QIAGEN CLC Genomics Workbench 20.0 (https://digitalinsights.qiagen.com)). The gene sequence information obtained from FGENESH was used as a reference and included in the alignment. Sequence polymorphisms were identified manually by comparing nucleotides of genes among all isolates.

3.2.4 Location of the *MAT1-2-1*

The location of *MAT1-2-1* of idiomorph *MAT1-2* was determined using sequence information from Warale (2015). This sequence was blasted against the *C. lentis* genome, and together with the genetic linkage map (Bhadauria *et al*., 2019), the location of *MAT1-2-1* was determined. Attempts were also made for the detection of the *MAT1-1* idiomorph. This was done using the

conserved alpha domain associated with this locus from various species such as *Neurospora crassa* and *Fusarium oxysporium*.

The positional conservation of the *MAT1* locus was confirmed by the flanking genes *SLA2*, *APN2, APC5* and *COX13* identified in various species and considered evidence for the presence of this locus (Dyer *et al*., 2016). These genes were blasted against the *C. lentis* genome.

3.3 Expression profiling of candidate genes in time course experiments

3.3.1 Isolates and mating experiments

Cross fertile isolates CT-30 (mIG-1) and CT-21 (mIG-2) identified by Armstrong-Cho *et al*. (2006) were used in the study. Isolates were cultured on OMA and incubated at 22°C with a 12 h photoperiod.

Mating experiments were conducted using sterilized lentil stems. Conidial suspensions were prepared by flooding 7 to 10-day old fungal cultures with sterile, distilled water, determining the spore concentration of suspensions with a haemocytometer and diluting them to obtain a final concentration of 2×10^5 conidia mL^{-1}. The sterilized plant stems were incubated for an hour in a mixture made up of a 5 mL conidial suspension of each of the two isolates to be mated. Five stems per Petri dish were placed on Whatman No. 1 filter paper overlaying agar media (prepared from deionised water), sealed with parafilm and incubated in the dark at 22°C and 70% relative humidity for 10 to 14 days (Armstrong-Cho and Banniza, 2006). The appearance of perithecia was monitored under a microscope at x1, x2 and x3 (Nikon SMZ-U Zoom and Nikon MKII Fibre Optic Lamp, Nikon, NY).

The isolates were also cultured individually and as a co-culture in three biological replicates for the harvest of mycelia. Mycelia of individual isolates was generated as described in Section 2.2.2. For co-culturing, the isolates CT-21 and CT-30 were first cultured separately on OMA and then conidial suspensions of each isolate was prepared with deionized sterile water and their spore concentration was determined with the haemocytometer. The suspensions of each isolate were diluted to 2×10^5 conidia mL^{-1} using sterile distilled water. Equal volumes of conidial suspensions from both isolates were mixed and spread on OMA plates. These plates were then incubated at 22°C with a 12 h photoperiod. After 7 to 10 days, the vegetative mycelia were

harvested from each replicate as described in Section 2.2.2. After lyophilisation, the mycelial samples were stored at -80ºC for RNA extraction.

3.3.2 Determination of different developmental stages of perithecia

Time-course experiments were conducted to evaluate the temporal dynamics of candidate gene expression. For this purpose, developmental stages of perithecia were distinguished on the basis of morphological characters including the shape and color of perithecia, as well as the formation of asci and ascospores inside the perithecium. To ensure consistency in assigning developing perithecia to a specific developmental stage, preliminary experiments were conducted to observe and describe perithecia on a regular basis for a period of approximately 30 days after first emergence of these fruiting structures. The location of a perithecium at the earliest stage of development on a lentil stem was identified, marked on the Petri dish, and examined and described daily using a dissecting microscope (Nikon SMZ-U Zoom with a Nikon MKII Fibre Optic Lamp, Nikon, NY). Images of the different developmental stages were taken and analysed using Zenn software (Carl Zeiss Vision Imaging Systems). The development of asci and ascospores was assessed with a Zeiss Axioskop 40 microscope (Zeiss Canada, Toronto, ON, Canada) after harvesting the perithecia of a particular stage from lentil stems using a fine, pointed scalpel and crushing them in a drop of deionized water on microscopic glass-slides. Slides were viewed under the microscope using x20, x40 and x60 objective lenses.

After identification and differentiation of defined developmental stages, mating experiments were set up on an ongoing basis, perithecia were harvested from every lentil stem, sorted and bulked based on their developmental stages and then flash-frozen in liquid nitrogen. Immediately after flash-freezing, the samples were stored at -80°C for use in RNA extraction. The perithecia of each developmental stage were harvested in three biological replicates, each of which consisted of perithecia harvested from mating experiments set up every week for one month. These mating experiments were performed in a similar way as described in 2.3.1 with inoculated lentil stems placed on Whatman No. 1 filter paper in approximately 30 agar media plates (5 stems per plate) every week. The perithecia of a particular stage were harvested from each plate at least 4 times until lentil stems were all black and covered with mycelia.

3.3.3 RNA extraction

3.3.3.1 RNA extraction from vegetative mycelium

The freeze-dried mycelial samples from each biological replicate (~ 100 mg) were ground using a battery-powered pestle mixer (Catalog No. 431-0100, VWR, part of Avantor), and homogenised for at least 1 min using a vortex mixer. Total RNA was extracted from the lysed mycelial samples using the RNeasy plant mini kit (QIAGEN, GmbH, Hilden, Germany) following the manufacturer's protocol with on-column DNase treatment. After homogenisation, the lysate was transferred to a QIAshredder spin column and centrifuged for 2 min at full speed. Subsequently, the flow-through was transferred to a new microcentrifuge tube without disturbing the pellet. Approximately 0.5 volume of 100% ethanol was added to the cleared lysate and thoroughly mixed by pipetting in and out. The lysate was centrifuged for 15 s at more than 10,000 rpm after transferring it to RNeasy spin column. The flow-through was discarded. The RNeasy spin column was then treated with 350 µL RW1 buffer, centrifuged for 15 s at more than 10,000 rpm and the flow-through discarded. After washing the spin column with RW1 buffer, it was treated with 80 µl of a DNase cocktail (10 µL of DNase1 stock solution + 70 µL RDD buffer) (QIAGEN RNase free DNase 1) for 15 min at room temperature. The spin column was then washed with 350 µL of RW1 buffer and centrifuged for 15 s at 12,000 rpm. The flow-through was discarded**.** The spin column membrane containing the sample was then washed with 500 µL of RPE buffer, centrifuged for 15 s at 12,000 rpm and the flow-through was discarded. This washing step with 500 µL of buffer RPE was repeated and the spin column was centrifuged for 2 min at 12,000 rpm. The flow-through was discarded. The RNeasy spin column was again centrifuged at 14,000 rpm for 2 min and the flow-through discarded. The RNeasy spin column was placed in a 1.5 ml micro-centrifuged tube (RNase free) and 40 µL of RNase free water was added directly to spin column. This was centrifuged for 1 min at 12,000 rpm and RNA eluted.

The quantity of RNA eluted per sample was determined using the A260: A280 ratio (~ 2.0) on a Nanodrop 8000 spectrophotometer (Thermo Fisher Scientific, Ottawa, ON). The integrity of RNA was checked by separating the 18S and 28S RNA on a 2.0% agarose gel (prepared in 1x Tris-acetate-EDTA (TAE) buffer (pH 8.0) and stained with Gel Red (Biotium Inc., CA)) and visualized under UV light using a Bio-Rad ChemiDoc^{TM} Gel Imaging System (Bio-Rad Laboratories, Inc., Canada).

3.3.3.2 RNA extraction from perithecia

The perithecia of a particular stage were harvested in pre-weighed 0.6 ml tubes, weighed and flash-frozen in liquid nitrogen immediately after harvesting and stored at -80°C for RNA extraction as described in Section 3.3.2. Approximately 20 mg of perithecia from each biological replicate of a particular development stage were required to yield sufficient amounts of RNA, which was achieved for two of four stages of development (Development stage 2 and stage 3). For each stage of development and for each individual biological replicate, perithecia were pooled in lysis matrix C 2 ML tubes (MP 116912; MP Biomedicals, Santa Ana, CA, United States) filled with 500 µL lysis buffer (RLC containing β-mercaptoethanol). The tissues were lysed in a tissue lyser for 1 min at maximum speed. After tissue lysis, the sample was again lysed with a battery-powered cordless pestle motor (Catalog No. 431-0100, VWR, part of Avantor) for a few seconds. Total RNA was extracted from the lysed mycelial samples using the RNeasy plant mini kit (QIAGEN GmbH, Hilden, Germany) with on-column DNase treatment following the manufacturer's protocol as described in Section 3.3.3.1 for three biological replicates of each stage. The quantity and quality of extracted RNA samples were measured using the A260:280 ratio (~2.0) on a Nanodrop ND8000 spectrophotometer (Thermo Fisher Scientific, Ottawa, ON).

3.3.4. cDNA synthesis and qPCR

The total RNA extracted from mycelial samples was diluted to 200 ng μL^{-1} using nuclease free ultra-pure water (Invitrogen Life Technologies, Carlsbad, CA, USA) for cDNA synthesis. As RNA yields from perithecia were significantly lower, cDNA was directly synthesized without diluting eluted RNA. The cDNA for all RNA samples (CT-21, CT-30, and co-cultures of CT-21 and CT-30, perithecia growth stages 2 and 3) was synthesized using high-capacity cDNA reverse transcription kits (Applied Biosystems™) following the manufacturer's protocol. The 20 µL reverse transcription reaction was prepared by adding 10 µL of diluted (mycelia) or undiluted (perithecia) RNA to 10 µL of the reagent mixture containing random primers, dNTPs, reverse transcriptase and nuclease free water. The synthesis was carried out in a Bio-Rad thermocycler (C1000 Touch™ thermal cycler, Bio-Rad Laboratories, Inc., Canada) and the reaction conditions consisted of 25°C for 10 min, 37°C for 120 min and 85°C for 5 min. The synthesized cDNA was stored at -20°C until used for qPCR. All cDNA samples were tested for gDNA contamination through PCR with exon-exon junction primers Cl_Nop2-qClMAT3, which

amplifies an amplicon of 266 bp for gDNA, whereas for cDNA it is 209 bp. PCR products of all samples were 209 bp, indicating no significant gDNA contamination.

All cDNA samples were diluted 5X and used as template for further gene expression studies through RT-qPCR. Primers spanning the intronic regions were designed from the known gene sequence information using PCR primer designing tool Primer3Plus Version 2.4.2 (Untergasser *et al*., 2012) (Table 3.2) for all but two genes *Cl_HECT* and *Cl_SART1*, which had no introns in their sequences so the primers were designed from exonic regions only. The amplification efficiency of each primer set was checked through RT-qPCR using QuantStudioTM3 System (Applied Biosystems Inc., Foster City, CA, USA) by using a cDNA dilution series, starting with 50 ng uL^{-1} and using a dilution factor of 2 (1:2 (V/V)). This was done four times in nuclease free ultra-pure water (Invitrogen Life Technologies, Carlsbad, CA, USA) for two biological replicates of mycelial sample CT-21. The qPCR reaction consisted of 1 µL cDNA, 5 µL PowerUp SYBER GREEN master mix (Applied biosystems), 0.4 µL each of forward and reverse primers (10 µM) and nuclease free ultra-pure water (Invitrogen Life Technologies, Carlsbad, CA, USA) bringing the final reaction volume to 10 µL. The primer efficiencies were calculated based on the slope of the standard curve using the equation: E = -1 + 10 (-1/slope). The acceptable amplification efficiency ranged from 90 to 120%. The house keeping gene β-tubulin (*Tub2*) was used as the reference gene (endogenous control) against all 11 genes. The primers for the gene were developed and provided by Dr. Jerlene Halliday (Research Officer, Pulse Crop Pathology Research Group, University of Saskatchewan, Saskatoon).

The RT-qPCR reaction for every gene was performed for three biological and three technical replicates in QuantStudio^{TM} 3 Real Time PCR System (Applied Biosystems Inc.) using a fast-run program. Each qPCR reaction consisted of 2 µL cDNA template, 5 µL SYBR green, 0.4 µL of each forward and reverse primers and 2.2 µL of nuclease-free water. The relative expression of the target genes was calculated as: $\Delta CT = \left(C_{T,Target} - C_{T,Reference}\right)$ as well as $2^{-\Delta C_T} = 2^{-(C_{T,Target} - C_{T,Reference})}$ (Livak and Schmittgen, 2001), where the housekeeping gene β-tubulin was used as the reference gene to normalize the transcript level of each gene. Expression data expressed as ΔCT were analyzed for each gene separately in SAS Version 9.4 (SAS Institute Inc, Cary, NC). Homogeneity of variances and normality of errors were confirmed with the Levene's and the Shapiro-Wilk test, respectively. The effect of developmental stage on gene expression

was analysed through ANOVA with the general linear model procedure of SAS and means were separated by Fisher's least significant difference at $P \leq 0.05$.

Chapter 4

4.0 Results

4.1 Identification of genes at QTL *qClMAT3*

A total of 106 genes, including 54 in the plus (+) strand and 52 in the minus (-) strand, were found at QTL *qClMAT3* using *Colletotrichum*-specific gene-finding parameters when the QTL sequence was subjected to *in silico* gene prediction using FGENESH (Table 4.1). Of 106 genes, 105 were complete genes and there were 16 genes with open reading frames. The coding sequences of genes were also determined and blasting (BLASTx) them against the NCBI non-redundant protein sequence database of the genera *Colletotrichum, Podospora* and *Neurospora* revealed a total of 105 homologs with known proteins (Tables 4.2, 4.3). One gene (*Cl_new*) present on the minus (-) strand had no homologous sequences in any of these species. Blasting these gene sequences against the *C. lentis* genome accessed on KnowPulse indicated that they are single copy genes.

Table 4.1 Structural features of genes underlying QTL *qClMAT3* of *Colletotrichum lentis* identified through *in silico* gene prediction using FGENESH. These genes were predicted using *Colletotrichum* specific gene-finding parameters

Features	**No. of genes**
Genes identified in + strand	54
Genes identified in – strand	52
Complete genes (transcription start site and poly A signal)	105
Genes with only transcription start site (no polyA signal)	1
Genes with open reading sequences (CDSo)	16

Sequence identity varied among homologue sequences of the different species. In the genus *Colletotrichum,* maximum homology for the genes underlying *qClMAT3* of *C. lentis* was identified in *C. higginsianum* followed by *C. graminicola*. Similarly, in the genera *Podospora* and *Neurospora*, maximum homology for the genes were found in the species *P. anserina*, *N. crassa* and *N. tetrasperma*. These species showed sequence identity of 50-100% and E-values of 0 for the majority of the gene sequences at *qClMAT3***.** The coding sequences were also blasted against the genus *Saccharomyces* but species showed homologies below 30% and with high E-values with the gene sequences at QTL *qClMAT3*.

Of 105 genes, six had protein domains that were common in both strands. These domains were an Aldo/keto reductase, an Rnase 3, a heat shock protein, an actin-cytoskeleton related protein, NADH-ubiquinone oxidoreductase and an integral membrane protein. The rest of the genes coded for different domains on both strands. Some of the domain-containing protein mRNAs include HECT, Denn, cytokinesis regulator mRNA, tRNA pseudouridine synthase mRNA, inner membrane protein and others. Hypothetical proteins were identified for 14 genes.

Table 4.2 Results from BLASTx search querying the coding sequences of genes identified in the plus (+) strand of QTL *qClMAT3* associated with mating incompatibility in *Colletotrichum lentis* against the genomes of species in *Colletotrichum*, *Neurospora* and *Podospora*.

Gene ID	**Domains**	**E-value**	**Percent identity**	**Accessions**
Cl_HECT	HECT domain	0	55	XP_011393585.1
Cl_Denn	Denn domain	0	99	XP_018154500.
Cl_GTPase	p-loop GTPase	0	70	XP_009852051.1
Cl_RNS3	RNAse III	0	98	XP_018154559.1
Cl_NADH	NADH-ubiquinone oxidoreductase 78 kDa subunit mRNA	7.00E-42	100	XP_018154525.1
Cl_SART1	SART-1 family protein mRNA	7.00E-139	51	XP_001909279.1
Cl_YdiU	YdiU domain-containing protein mRNA	0	98	XP_018154563.1
Cl_FUT3	Alpha (1,3)-fucosyltransferase 3 mRNA	0	88	XP_018154539.1
Cl_HP1	Hypothetical protein mRNA	0	97	XP_018154538.1
Cl_ITP	Intracellular protein transport protein mRNA	0	55	VBB77382.1
Cl_PUS	tRNA pseudouridine synthase mRNA	0	98	XP_018154552.1
Cl_Nop2	NOL1/NOP2/sun family protein mRNA	0	97	XP_018154566.1
Cl_Hsp60	Heat shock protein 60 mRNA	0	100	XP_018154570.
Cl_Duf974	Duf974 domain containing protein mRNA	0	96	XP_018154501.1
Cl_Lin1	Lin1 family protein mRNA	0	94	XP_018154542.1
Cl_BHLH	BHLH family transcription factor mRNA	9.00E-24	34	XP_001909620.1
Cl_TASP	Threonine aspartase mRNA	9.00E-170	57	XP_001909552.1
Cl_ASNS	Asparagine synthase mRNA	0	99	XP_018154474.1
Cl_HP2	Hypothetical protein mRNA	0.0004	32	XP_001909595.1
Cl_ADH	Alcohol dehydrogenase 1 mRNA	0	100	XP_018154458.1
Cl_Pumilio	Pumilio-family rna binding repeat protein mRNA	0	96	XP_018154557.1
Cl_BING4CT	BING4CT domain-containing protein mRNA	0	71	XP_009851929.1
Cl_Epi	NAD dependent epimerase/dehydratase mRNA	7.00E-177	77	XP_018154536.1
Cl_COX18	Inner membrane protein COX18 mRNA	1.00E-33	76	XP_009852268.1
Cl_Duf974	Duf974 domain-containing protein mRNA	0	95	XM_018305118.21
Cl_HP3	hypothetical protein mRNA	7E-10	38	XP_001909590.1
Cl_Cro1	Actin cytoskeleton organization protein mRNA	0	65	VBB77716.1
Cl_HP4	hypothetical protein mRNA	0	97	XP_018154487.1
Cl_HP5	hypothetical protein mRNA	5.00E-20	41	XP_001909519.1
Cl_AKR	Aldo/keto reductase mRNA	0	98	XP_018154564.1
Cl_Cabeza	Rna-binding protein cabeza-like protein mRNA	0	82	XP_018154572.1
Cl_Hsp60	Heat shock protein 60 mRNA, chaperonin GroL partial mRNA	0	98	XP_018154570.
Cl_Cofilin	cofilin/tropomyosin-type actin-binding protein mRNA	4.00E-106	99	TIC97153.1
Cl_HP6	Hypothetical protein	7.00E-135	99	XP_018154568.1
Cl_MA	Mitotic apparatus protein p62-like protein mRNA,hypothetical protein	3.00E-145	87	XP_018154565.1
Cl_Mrp	Mitochondral 37s ribosomal protein mRNA	8.00E-57	100	XP_018154558.1
Cl_ Pol III	RNA polymerase III subunit	1.00E-90	98	XP_018154554.1
Cl_GPI	GPI-GlcNAc transferase complex partial mRNA	0.00E+00	98	XP_018154552.1
Cl_PAN1	Actin cytoskeleton-regulatory complex protein PAN1-like protein 1	2.00E-60	90	XP_018154548.1
Cl_HP7	Hypothetical protein	0.00E+00	97	XP_018154546.1
Cl_RRM	RNA recognition motif containing protein	5e-108	98	XP_018154526.1
Cl_COX6B	Cytochrome c oxidase subunit 6B mRNA	9.00E-40	100	XP_018154524.1
Cl_Plectin	Plectin/S10 domain-containing protein mRNA	2.00E-69	100	XP_018154522.1
Cl_GroES	Alcohol dehydrogenase GroES-like domain-containing protein mRNA	6.80E+00	54	XP_001908855.1
Cl_HP8	Hypothetical protein	0	98	XP_018154519.1
Cl_IPP1	Type1 protein phosphatase regulator/inhibitor	6.00E-21	82	XP_001909452.1
Cl_OTC	Ornithine carbamoyltransferase mRNA	0	97	XP_018154504.1
Cl_KH	KH domain-containing protein mRNA	0.00E+00	71	XP_011393474.1
Cl_KOW	KOW domain-containing protein mRNA	0	94	XM_018305068.1
Cl_Hpt	Hpt domain-containing protein mRNA	7.00E-54	55	XP_001909522.1
Cl_HXT	hexose transporter	5.00E-40	69	XP_001909593.1
Cl_PSA2	Proteasome subunit alpha type-2 mRNA	6.00E-146	86	XP_001909600.1
Cl_ GGT	Gamma-glutamyltranspeptidase mRNA	0.00E+00	97	TIC96988.1
Cl_SNARE	SNARE domain-containing protein mRNA	1.00E-33	76	XP_009852268.1

Table 4.3 Results from BLASTx each querying the coding sequences of the genes identified in the minus (-) strand of QTL *qClMAT3* associated with mating incompatibility in *Colletotrichum lentis* against the genome of species in *Colletotrichum*, *Neurospora* and *Podospora.*

Gene ID	**Domain**	**E-value**	**Percent identity**	**Accessions**
Cl_Ricin	Ricin B lectin mRNA	8.00E-101	83	EXF86655.1
Cl_HP9	Hypothetical protein	7E-135	99	XP_018154568.1
Cl_Byr4	Cytokinesis inhibitor byr4	0	97	XP_018154566.1
Cl_AKR	Aldo/keto reductase	0	98	XP_018154564.1
Cl_NADH	NADH:ubiquinone oxidoreductase 78	0	85	XP_957188.3
Cl_ACBP	Acyl CoA binding protein mRNA	0	99	XP_018154562.1
Cl_ACBP	Acyl CoA binding protein mRNA	1E-174	96	XP_018154560.1
Cl_RNS3	RNase3 domain-containing protein	6E-89	47	XP_011393503.1
Cl_HP10	hypothetical protein partial mRNA	1E-140	87	TKW53373.1
Cl_RPPK	ribose-phosphate pyrophosphokinase	0	97	CCF37957.1
Cl_IDC1	IDC1 protein	0	90	KZL82142.1
Cl_HP11	Hypothetical protein	0	97	XP_018154549.1
Cl_UFD	ubiquitin fusion degradation protein	0	99	XP_018154543.1
Cl_RuvBL2	RuvB-like helicase 2	0	100	XP_018154541.1
Cl_GAPDH	GAPDH	0	100	KZL77452.1
Cl_PfpI	DJ-1/PfpI family protein	6E-151	93	XP_018154535.1
Cl_IMP	Integral membrane protein	0	98	XP_018154534.1
Cl_C6TF	C6 transcription factor	8E-58	78	XP_018154531.1
Cl_CpB	Capsule polysaccharide biosynthesis protein	0	96	XP_018154530.1
Cl_Ankyrin	Ankyrin repeat protein	3E-41	68	KZL77493.1
Cl_Sad1	hypothetical protein /sad1 interacting protein	1E-152	95	XP_018154528.1
Cl_RRM	RNA recognition domain-containing protein mRNA	0	97	XP_018154527.1
CL_NADH	NADH:ubiquinone oxidoreductase 78	7E-42	100	XP_018154525.1
Cl_SDE2	Telomere maintenance protein SDE2	7E-176	98	XP_018154523.1
Cl_MRP49	ribosomal protein MRP49	9E-103	79	XP_018154517.1
Cl_AARP2CN	AARP2CN	0	99	XP_018154513.1
Cl_PtRH2	peptidyl-tRNA hydrolase 2	9E-119	99	TIC97491.1
Cl_GAM	glucoamylase	0	99	XP_018154509.1
Cl_PrRP	Pre-rRNA processing protein	0	98	XP_018154505.1
Cl_mtALS	mitochondrial acetolactate synthase small subunit precursor	0	99	TKW59407.1
Cl_ASNAP	alpha-soluble NSF attachment protein	0	98	XP_018154499.1
Cl_Duf726	DUF726 domain-containing protein	0	94	XP_018154497.1
Cl_P450	Benzoate 4-monooxygenase cytochrome P450	2E-111	77	XP_018154496.1
Cl_ PolI	DNA polymerase iota	0	93	XP_018154495.1
Cl_RhoA	GTP-binding protein rhoA	8E-135	100	ELA37724.1
Cl_IK13	IKI3 family protein	0	97	TIC97103.1
Cl_Hsp70	Hsp70 protein	0	99	XP_018154484.1
Cl_MAK	Serione -theorine protein kinase mak	0	99	TIC97464.1
Cl_PDK	pyruvate dehydrogenase kinase	0	99	XP_018154480.1
Cl_SPL	Sphingosine-1-phosphate lyase	0	98	XP_018154471.1
Cl_ManT	alpha-1,6-mannosyltransferase subunit	0	94	XP_018154470.1
CL_RFD	Ring finger domain protein	0	95	XP_018154469.1
Cl_WD40	WD40 repeat-like protein	0	98	XP_018154461.1
Cl_P450	Cytochrome P450 52A4	0	97	TIC97398.1
Cl_TFb4	Transcription factor tfb4	0	99	XP_018154456.1
Cl_Trm	tRNA methyltransferase	0	94	XP_018154454.1
Cl_Alg11	Alpha-1,2-mannosyltransferase alg-11	0	98	XP_018154452.1
Cl_Actin	Actin	0	100	XP_018154458.1
Cl_HP12	Hypothetical protein	0	85	XP_018154472.1
Cl_HP13	Hypothetical protein	2E-166	89	CCF46616.1
Cl_HP14	Hypothetical protein	0	94	XP_018154465.1
Cl_new	No homology	-	-	-

4.2 Identification of candidate genes for mating incompatibility

4.2.1 Identification of candidate genes for mating incompatibility

The functional role of genes at QTL *qClMAT3* was inferred by comparing their sequences with the nucleotide or protein sequences of the closely related species as well as model organisms. Twenty-eight predicted genes were identified as homologs of interest as they may be associated with the mating pathway (Table 4.4). Most of these proteins were regulatory proteins, and seven domain-containing proteins were identified that are involved in cell cycle regulation, e.g. a protein with a HECT (Homologous to the E6-AP Carboxyl Terminus) domain. HECT domain-containing proteins are mostly ubiquitin-protein ligases, which cause the ubiquination of cdc25 proteins, which, in turn, control the cell cycle. There were also two protein-encoding genes, *ICD 1* (inhibitor of cell death) and actin-cytoskeleton organization (*CRO*), which play a direct role in sexual reproduction of ascomycetes. The role of *IDC1* as a part of a MAP (mitogen-activated protein) kinase signalling cascade and in sexual reproduction has been reported in *P. anserina* (Jamet-Vierny *et al*., 2007). Likewise, the CRO1 protein was found to play a role in dikaryotic cell divisions and ascospore development. Other domains such as Denn domain, DUF974 domain, Hpt domain, serine-theorine protein kinase and intracellular membrane protein domains are a part of signalling pathways. Most of these domains are related to G-protein-coupled receptors that are involved in pheromone signalling pathways and hence the initiation of sexual reproduction (Xue *et al*., 2008). Likewise, domains such as RNaseIII are involved in silencing events, and others such as mitotic apparatus protein are recombination or double strand break (DSB) repair proteins. Some of the domain-containing proteins have conserved regions of tryptophan (W) and aspartic acid (D) dipeptide (WD-40) repeats, which are known to play a role in vegetative incompatibility.

4.2.2 Status of candidate genes in mating incompatibility groups of *Colletotrichum lentis*

4.2.2.1 Presence/ absence of genes identified in isolates of mIG-1 and mIG-2

Screening for the presence or absence of 28 genes in 20 *C. lentis* isolates by PCR revealed that 11 genes (*Cl_HP9 _NOP2, _3IMP, _BHLH, _HECT, _SART1, _MAK, _HP3, _RNS3, _WD40, _Cro1*) were present in all isolates (Table 4.5). The remaining 17 genes (*Cl_Byr4, _HP1,*

Table 4.4 *Colletotrichum lentis g*enes at QTL *qClMAT3* coding for protein domains potentially involved in mating incompatibility, and their functions inferred from the literature and through protein family and domain databases such as UniProtKB and InterPro.

Gene ID	Domain- containing proteins	Functions	References
Cl_HP9	Hypothetical protein	Blt 1, cell cycle control	Moseley *et al*., 2009
Cl_Nop2	NOL1/NOP2	Cell cycle	Beus *et al*., 1994
Cl_Byr4	Cytokinesis inhibitor byr4	Cell division control	Song *et al*., 1996
Cl_MA	Mitotic apparatus protein	DNA binding and DSB repair	NCBI
Cl_RNS3	RNAseIII	Involved in RNAi or gene silencing	Carmell *et al*.,2004
Cl_IDC1	IDC1 protein	IDC gene involved in sexual reproduction	Gautier *et al*., 2018
Cl_PAN1	PAN1 like protein	Cell division	NCBI
Cl_HP7	Hypothetical protein	Transcription factor	NCBI
Cl_Pfp1	DJ-1/Pfpl family protein	Oxidative stress response	NCBI
Cl_IMP	Inner membrane protein	Specificity of intracellular membrane fusion processes fat storage and longevity	Exil *et al*., 2010
Cl_BHLH	BHLH family transcription factor	Vegetative incompatibility protein	Espagne and Barreau*,* 2002
Cl_HECT	HECT domain	Ubiquination of CDC25 (involved in various phases of cell cycle)	Rotin and Kumar., 2009
Cl_HP8	p-loop GTPase	G protein alpha subunit	NCBI
Cl_IPP1	Type 1protein phosphatase regulator	Required for mitosis/serine/threonine phosphatase inhibitor activity	NCBI
Cl_BING4CT	BING4CT	Vegetative incompatibility protein	NCBI
Cl_SART1	SART-1 (Homolgue of snu66 of yeast)	Cell cycle arrest,pre-mRNA splicing	Wilkinson *et al.,* 2004
Cl_KH	KH domain	RNA binding protein involved in phermone signalling	NCBI
Cl_Duf974	Duf974	Trafficking protein	Ramirez-Peinado *et al.*, 2017
Cl_Denn	Denn domain	Regulation of MAPK signalling pathway	Marat *et al*., 2011
Cl_Ik13	Ik13	Transcriptional factor (Elongator)	NCBI
Cl_Hpt	Hpt domain	Phosphorelay signal system that is involved in suppression of mating when sufficient nutrients are available	Maksimov *et al*.,2016
Cl_MAK	Serine-theorine protein kinase	MAPK signaling	NCBI
Cl_HP3	Hypothetical protein	Silencing (Stc1domain-containing protein)	Bayne *et al*., 2010
Cl_RNS3	Rnase 3	HET /Dicer protein	Carmell *et al*.,2004
Cl_IMC	Intracellular	Signaling pathway protein	Nelson, 2003
Cl_WD-40	WD-40 repeat like protein	Vegetative incompatibility protein	NCBI
Cl_Cro1	Actin cytoskeleton organisation protein	CRO1 protein involved in sexual reproduction	Berteaux-Lecellier *et al*., 1998
Cl_SNARE	SNARE domain	Specificity to the intracellular membrane fusion processes	NCBI

_RNSIII, _IDC1, _PAN1, _HP3, _IMP, _HP8, _IPP1, _BING4T, _KH, _Duf974, _Denn, _IK13, Hpt, _HP2, _SNARE) were present or absent in *C. lentis* isolates of both mating incompatibility groups, indicating that they are not associated with mating incompatibility. For instance, *HP7* was absent in isolates of mIG-1 (GT-99, GT-119, GT-164) as well as isolates of mIG-2 (GT-75, GT-125 and GT-146).

In some isolates, such as GT-75, GT-125 and GT-164 almost 50% of these candidate genes were absent. In contrast, there were also some isolates such as GT-29, GT-64, GT-96, GT-105 and GT-120, in which most of the genes were present. The absence of these genes was confirmed by performing the PCR experiments in two biological replicates.

All 28 genes were present in CT-30. However, there were some genes (*Cl_Pfp1*, *Cl_MA, Cl_Hpt, Cl_KH* and *Cl_PAN1)* which were absent in the other parental isolate CT-21.

Table 4.5 Presence (+) and absence (-) of 28 genes located at QTL *qClMAT3* coding for protein domains potentially involved in mating incompatibility in 20 *Colletotrichum lentis* isolates based on PCR with gene-specific primers

Isolates	mIG-1											mIG-2										
Genes	CT-30	GT-2	GT-12	GT-29	GT-30	GT-64	GT-99	GT-105	GT-119	GT-120	GT-164	CT-21	GT-7	GT-52	GT-69	GT-75	GT-96	GT-111	GT-125	GT-135	GT-138	GT-146
Cl_HP9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_Nop2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
Cl_Byr4	+	+	-	+	+	+	+	+	+	+	+	+	-	+	-	+	-	+	+	+	+	-
Cl_MA	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-	+	+	+	-	+	+	+
Cl_RNSIII	+	+	+	+	+	+	-	+	+	+	-	+	+	+	+	-	+	+	-	-	+	-
Cl_IDC1	+	+	+	+	+	+	-	+	+	+	-	+	+	+	+	-	+	+	-	-	+	-
Cl_PAN1	+	+	+	+	+	-	+	+	+	+	+	-	+	+	+	-	+	+	+	+	-	+
Cl_HP7	+	+	+	+	+	+	-	+	-	+	-	+	+	+	+	-	+	+	-	+	+	-
Cl_Pfp1	+	-	+	+	-	+	+	+	+	+	-	-	+	-	+	-	+	+	+	+	-	+
Cl_IMP.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_BHLH	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_HECT	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_GTPase	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-
Cl_IPP1	+	+	+	+	+	+	-	+	+	+	-	+	+	+	+	-	+	+	+	+	+	-
Cl_BING4CT	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+
Cl_SART1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_KH	+	+	+	+	-	+	+	+	+	+	-	-	+	+	+	-	+	+	-	+	+	+
Cl_Duf974	+	+	-	+	+	+	+	-	+	+	-	+	+	+	-	-	+	-	+	+	-	-
Cl_Denn	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+
Cl_Ik13	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	-	+	+	-	-
Cl_Hpt	+	+	+	+	-	+	+	+	+	+	+	-	+	-	+	-	+	+	+	+	+	+
Cl_MAK	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_HP3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_RNS 3	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	-	+	-
Cl_IMP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_WD40	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_Cro1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cl_SNARE	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+

4.2.2.2 Sequence polymorphisms of candidate genes in isolates of mIG-1 and mIG-2 of *Colletotrichum lentis.*

Candidate gene sequence comparison among isolates revealed approximately 90% sequence identity. Polymorphisms detected among isolates were not correlated with either mIG-1 or mIG-2, and involved both transitions as well as transversions. However, there were more sequences showing transitions (C→T) than transversions. The highest number of transitions occurred in sequences of *Cl_Nop2* among isolates. There were also gene sequences showing both transitions and transversions, e.g in sequences of *Cl_BING4CT* and *Cl_IPP1*. Sequences of *Cl_Cro1* showed very few polymorphisms among isolates (Appendix 1-11).

4.2.2.3 Location of *MAT1-2-1*

Using the sequence information from Warale (2015) and genetic map information available from the *C. lentis* genome (Bhadauria *et al*., 2019), the *MAT1-2-1* location was identified on scaffold 1: 4514190- 4515048 of the minus (-) strand of CT-30. This scaffold is part of chromosome 9 and lies within the extended SNP marker position 4640612-1122745 with original marker position scaffold1p: 4326227- 1134067. The sequence present at this position was confirmed to be *MAT1-2-1* by blasting the coding sequence of this gene with the *C. lentis* genome accessed in KnowPulse (http://knowpulse2.usask.ca/portal/node/1) as well as on NCBI against sequences of species in the genus *Colletotrichum*. This gene showed sequence homology with the HMG box domain containing protein of *Colletotrichum higgisianum* and *Colletotrichum graminicola*. The gene *MAT1-2-1* has been structurally annotated as *g8726* in the *C. lentis* genome available through KnowPulse with a single copy in the genome. This locus was further confirmed to be the *MAT1* locus based on the positional conservation of various genes flanking this locus, such as *SLA1* downstream to the *MAT1* locus, and *APN2, APC5* and *COX13* upstream of the locus, and *APC5* in-between *APN2* and *COX13*. All attempts to find *MAT1-1* failed in this study indicating that it is absent from the *C. lentis* genome.

4.3 Expression profiling of candidate genes in time course experiments

4.3.1 Determination of the developmental stages of perithecia

On the basis of visual differences in the development of perithecia, four developmental stages were differentiated for time-course experiments to evaluate the temporal dynamics of candidate gene expression (Figures 4.1, 4.2):

Stage 1: During this stage, the perithecia appeared as transparent, flask-shaped fruiting structures. The asci in the perithecia started to form and were immature.

Stage 2: During this stage, the perithecia started to turn brown-blackish from the base. Based on microscopic examination of crushed perithecia, the asci in the perithecia started maturing and formed ascospores but less than eight ascospores were present per ascus.

Stage 3: During this stage, perithecia were identified as completely black-coloured structures, which had fully developed asci and ascospores and at this stage asci were emerging from the ostiole.

Stage 4: During this stage, perithecia had either burst and released all ascospores or contained only a few remaining asci, but most of the perithecia harvested had no asci. The perithecia were identified as black structures that were open on one side, and appeared as complete structures from one side, but were in fact empty structures with only the perithecial envelope remaining.

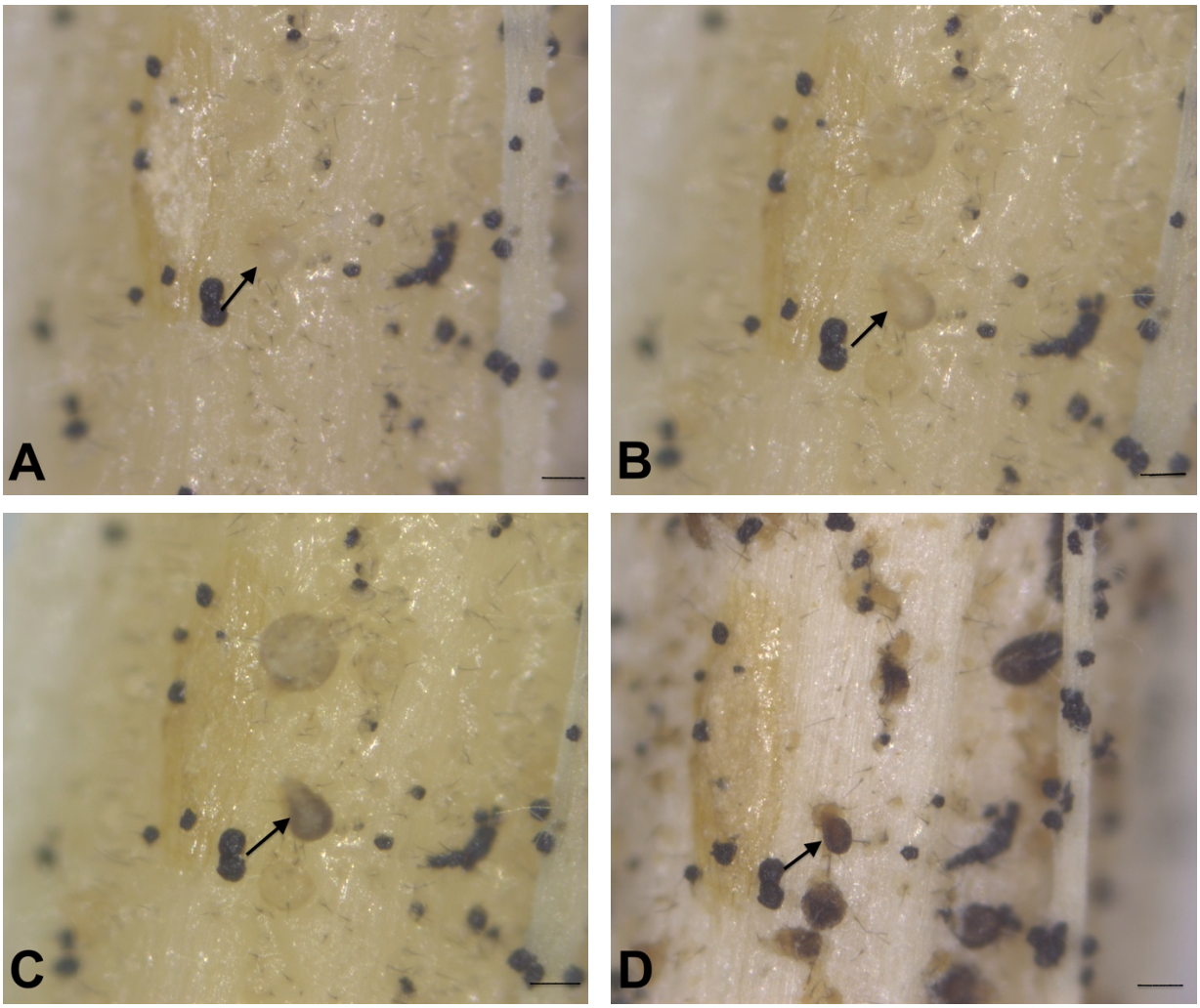


Fig 4.1. Photomicrographs of the developmental stages of perithecia (arrows) of *Colletotrichum lentis*. **A**. Developmental stage 1 perithecia transparent. **B**. Developmental stage 2 perithecia started to turn black from the base. **C**. Developmental stage 3 perithecia became black-brown. **D**. Development stage 4 perithecia burst from one side. (Scale bar=100µm)

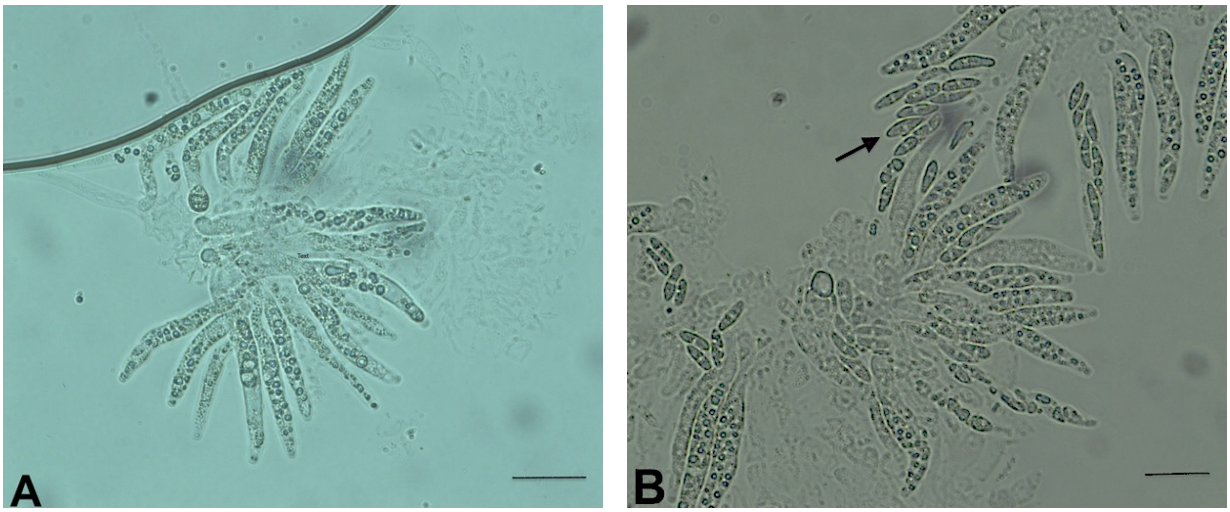


Fig.4.2. Photomicrographs of asci and ascospores inside perithecia of *Colletotrichum lentis* at stages 2 and 3**. A**. Developmental stage 2 when the asci had developed. **B**. Developmental stage 3 when the ascospores (arrow) had developed in most asci. (Scale bar=100µm)

4.3.2 Expression of candidate genes in co-culture of CT-21 and CT-30 and in developmental stages of perithecia.

To analyze changes in expression level of various genes during vegetative and reproductive stages, the variation in gene expression was compared between co-cultures, developmental stages 2 and stage 3 of perithecia and individual mycelia of isolates of CT-30 and CT-21. Comparison of the relative gene expression (ΔCT) of genes in mycelia of isolates CT-21 and CT-30, co-cultures and perithecial stages 2 and 3 revealed that there were no significant differences in expression levels of *Cl_Byr4* and *Cl_HECT*.

Cl_Nop2 was highly expressed in individual mycelium of CT-30 compared to CT-21, whereas for *Cl_IDC1* it was the reverse. Similar to mycelium of CT-21, gene expression of *Cl_IDC1* was upregulated in the co-culture and in perithecial stage 3, whereas in stage 2 it was down-regulated similar to CT-30. *Cl_HP9*, which encodes for a cell-cycle protein (Table 4.4) and *Cl_BING4CT*, which encodes a vegetative incompatibility protein were downregulated in the co-culture compared to all other stages. In contrast, *Cl_SART1*, *Cl_KH*, *Cl_Hpt* and *Cl_HP3* were all upregulated in the co-culture compared to individual mycelia, stage 2 perithecia, and in the case of *Cl_Hpt* and *Cl_KH* also in comparison to stage 3 perithecia. There was one gene, *Cl_Cro1*, whose expression level at stage 2 was significantly downregulated compared to all other stages.

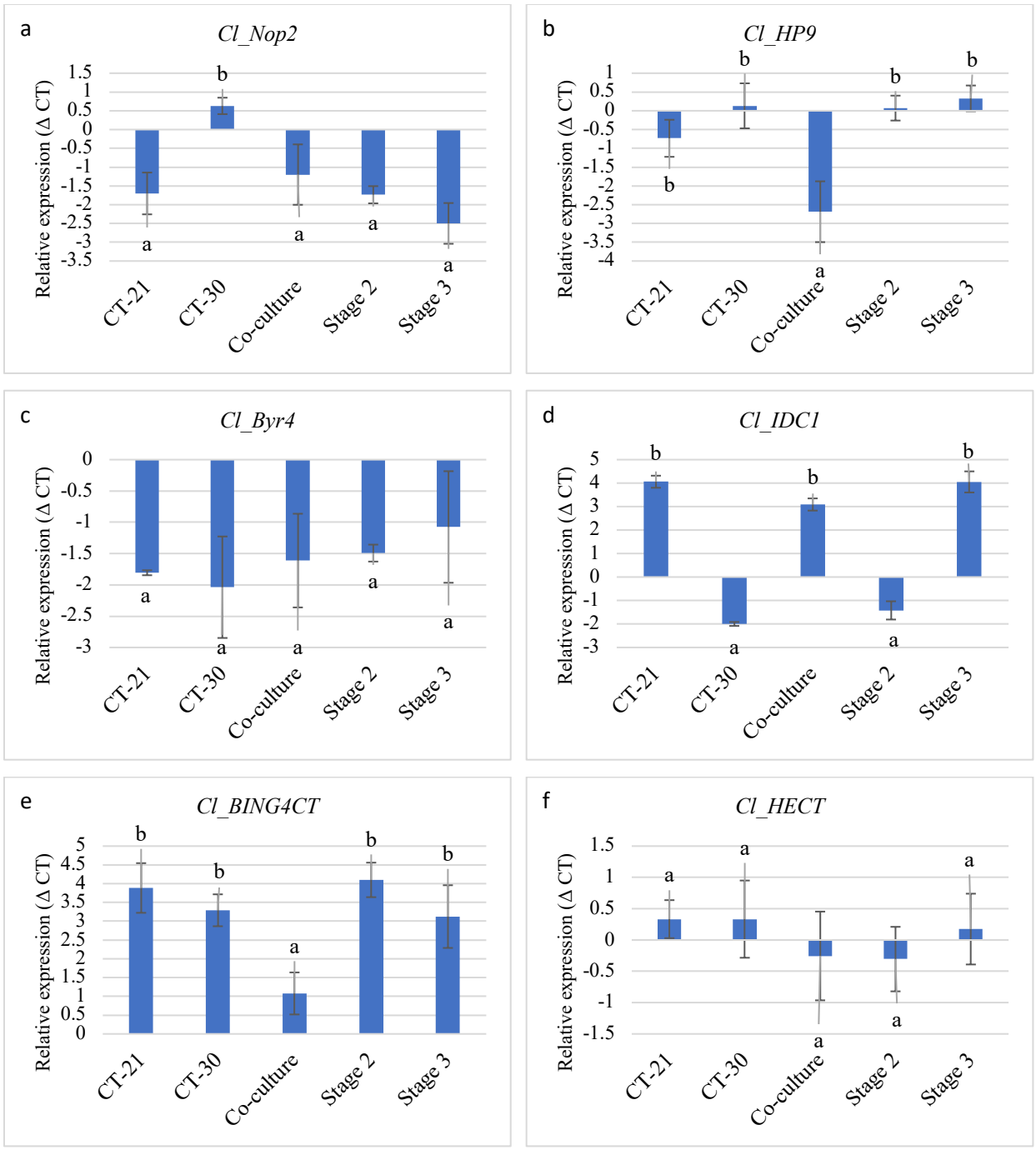
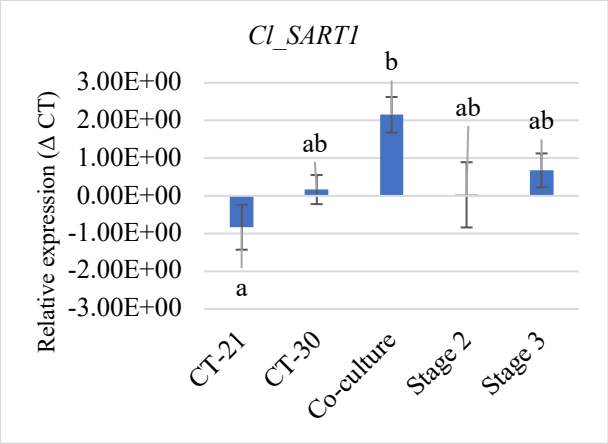
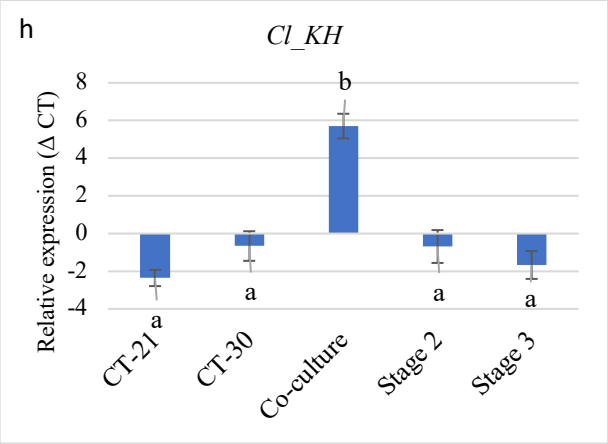
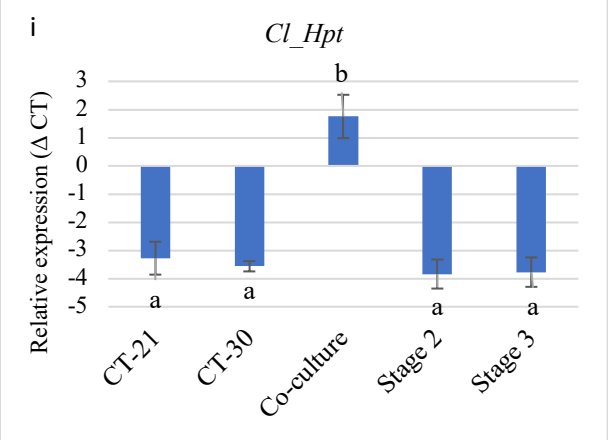
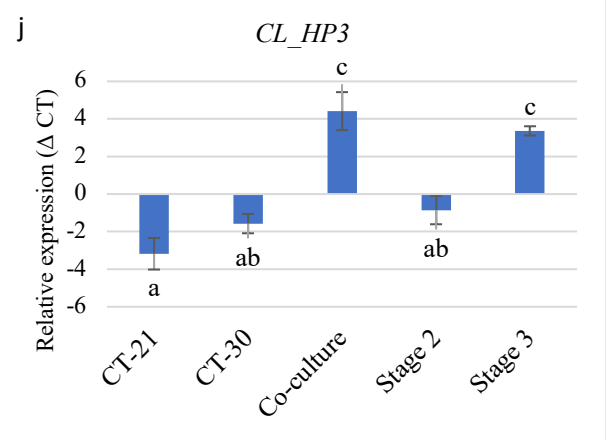


Fig.4.3 a-k. Relative expression level calculated as Δ*CT* of 11 candidate genes located at QTL *qClMAT3* coding for protein domains potentially involved in mating incompatibility in *Colletotrichum lentis* at various vegetative and reproductive stages**.** Bars with the same letter are not significantly different from each other (***Cl_Nop2***: *P* = 0.0167; ***Cl_HP9***: P = 0.0154; ***Cl_IDC1***: P <0.0001; ***Cl_BING4CT***: P = 0.0367; ***Cl_KH***: P< 0.0001; ***Cl_Hpt***: P < 0.0001; ***Cl_HP3***: P < 0.0001; ***Cl_Cro1***: P =0.0132.









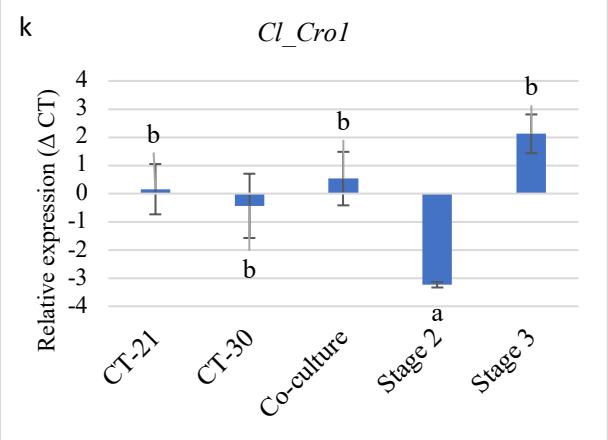


Fig.4.3 a-k. Continued

Chapter 5

5.0 Discussion

The overall objective of the current study was to understand the mating system of *C. lentis* in depth, mainly to unravel the genetic basis of mating incompatibility. Although the occurrence of sexual reproduction in most *Colletotrichum* species including *C. lentis* is very rare or presumed absent in field conditions (Menat *et al*., 2012), the possibility of mating under field conditions carries the risk of genetic variability occurring as a result of sexual recombination. This could lead to the development of new races and breakdown of resistance of currently available anthracnose resistant lentil varieties. Consequently, the study of the mating system is important to understand the sexual behaviour of *C. lentis* and specifically the molecular mechanisms regulating the first step of mating, i.e. compatibility between the parental isolates. *Colletotrichum lentis* appears to follow a bipolar mating system based on the presence of two mating incompatibility groups (Armstrong-Cho and Banniza, 2006), but the *MAT1* locus does not regulate mating incompatibility (Meant *et al*., 2012). Based on bulk segregant analysis, major QTL *qClMAT3* located on LG3 could be associated with mating incompatibility (Bhadauria *et al*., Dept. of Plant Sciences, University of Saskatchewan, unpublished). In the current study, genes other than those associated with *MAT1-2* and *MAT1-1* idiomorphs were identified at QTL *qClMAT3* that probably play a role in regulating mating incompatibility. *MAT1-2-1* was not only absent at this QTL, but was located on another chromosome, chromosome 9 of the *C. lentis* genome. None of the candidate genes were present or absent in one, but not in the other mIG, and none showed the same sequence polymorphism in all isolates of one mIGs. However, they displayed interesting expression patterns at various vegetative and reproductive stages, especially in co-cultures, indicating these genes as most promising candidates regulating mating incompatibility.

Among 28 genes associated with cell cycling, mating or sexual reproduction, 11 most likely candidate genes associated with mating incompatibility were selected based on their functional role in other organisms. Two genes, *Cl_Byr4* and *Cl_ HECT,* showed no differential expression in individual or combined vegetative mycelia or in perithecial developmental stages 2 and 3. *Cl_Byr4* encodes for cytokinesis inhibitor protein and *Cl_HECT* encodes for cell cycle protein

involved in ubiquination of CDC25. Their consistent expression pattern during all developmental stages indicated that they are most likely not involved in regulation of mating incompatibility.

In contrast, *Cl_Nop2* and *Cl_IDC1* showed differential expression in the vegetative mycelium of CT-21 and CT-30. *Cl_Nop2* was highly expressed in CT-30 compared to CT-21, whereas *Cl_IDC1* was highly expressed in CT-21 compared to CT-30. *Cl_Nop2* encodes for an RNA methyltransferase and plays a role in cell proliferation and viability. Expression of this gene at various vegetative or reproductive stages has not been documented in model ascomycetes species. *Cl_IDC1* encodes for a pezizomycotina specific *IDC1* (inhibitor of cell death) protein and is homologous to *IDC1* of *Podospora anserina* and *Ham-5* of *N. crassa*. In *P. anserina*, *IDC1* is involved in the localisation of MAP kinase 1 (Mpk1) to the nucleus during chemotropic interactions and is expressed in the female partner leading to the development of protoperithecia from ascogonium, and finally to perithecia (Jamet-Vierny *et al*., 2007). In *N. crassa*, classical genetic studies revealed that the homologue of *IDC1*, *Ham5* (hyphal anastomosis mutant gene 5) is required for hyphal fusion as it is involved in localisation of the MAK-2 signalling cascade to the correct cellular location during chemotropic interaction (Jonkers *et al*., 2014). Similar expression patterns of *Cl_IDC1* were observed using RT-qPCR where this gene was significantly expressed in one isolate, CT-21, but not in its mating partner, CT-30. These expression patterns need to be further validated with other compatible isolates belonging to mIG-1and mIG-2 of *C. lentis* to confirm that *Cl_IDC1* is consistently upregulated in isolates of mIG-1 and downregulated in isolates of mIG-2. Interestingly, this gene was also upregulated in the vegetative co-cultures of both isolates and in perithecial stage 3, but not stage 2. This suggests that in *C. lentis* this gene has no role in the dikaryon formation that occurs in perithecial stage 2. Instead, it may be involved in the signaling pathway during chemotropic interaction and in the development of protoperithecia.

In the co-culture of vegetative mycelium of CT-21 and CT-30, *Cl_KH* and *Cl_Hpt* were significantly upregulated compared to all other developmental stages. *Cl_KH* encodes a KH domain encoding protein. Pheromones and their cognate receptors play a vital role in the regulation of mating, particularly during the initial stages. When the pheromones of *S. cereviseae* bind to the receptors (hetrotrimeric G- proteins Gαβγ) present on the cell surface of the opposite mating partner, the GPA-1 subunit of the Gα gets activated along with a Gβγ dimer (Guo *et al*., 2003). This activation leads to the dissociation of Gα from the Gβγ, the hydrolyzation of GTP to

GDP takes place on Gα and mating signaling is amplified through Gβγ release. This activation initiates a mitogen activated protein kinase (MAPK) signaling module, which regulates the expression of mating-specific genes. In addition to signaling via Gβγ, pheromone signaling through mating factor-activated G-protein α-subunit (Gpa1) effector protein Scp160 (Saccharomyces protein controlling ploidy,160kDA) has also been demonstrated by Guo *et al*. (2003). This RNA binding KH effector protein plays a role in the localization of mRNA and facilitates the accumulation of components involved in cell polarization and mating at the site of polarized membrane extensions known as shmoos. When haploid yeast cells (*MAT*a or α) were treated with the opposite mating factor, shmoos were formed in the direction of the pheromone gradient. Gelin-Licht *et al*. (2012) determined the role of this protein during mating signaling and in mating efficiency. The mislocalisation of mitogen-activated kinase FUS3 proteins in Scp160 mutants indicated its role in localisation of proteins at the shmoo tip. It was also found that SRO7 and SEC3 POL mRNAs were significantly enriched in α-mating factor treated *MATa* cells and activity of Scp160 also increased in treated cells as compared to untreated cells indicating its role in mRNA trafficking during pheromone signaling. This effector protein was also found to affect mating efficiency through RNA binding function as the mating efficiency was inhibited in heterozygous crosses between cells with mutations in the KH domain of Scp160 and wild type cells (Gelin-Licht *et al*., 2012). In the present study, *Cl_KH* encoding for a KH domain encoding protein was significantly upregulated when two mating partners were grown together as compared to mycelia of individual isolates or developing fruiting structures. Based on the function of homologues in model species, this observation could indicate a possible involvement of *Cl_KH* in the pheromone signaling pathway.

Cl_Hpt encodes a histidine phosphotransfer (HPt), which was shown to be essential for viability of spores in *N. crassa*, *S. cereviseae*, *Cryptococcus neoformans* and *Aspergillus nidulans* and has been reported to play a critical role in asexual reproduction through its expression in response to light in *M. oryzae* (Mohanan *et al*., 2017). It is part of a two component phosphorelay system that regulates various cellular functions such as oxidative and osmotic stress responses, differentiation, sexual and asexual development. This HPt together with histidine kinases (HK) are responsible for suppression of stress responsive high osmolarity glycerol (HOG) pathway in most fungal species. Most fungi respond to environmental stresses through a multicomponent phosphorelay signal system consisting of two upstream branches. One branch consists of a three-

component phosphorelay system: hybrid histidine kinases (HHKs), histidine containing phosphotransfer (HPt) protein and response regulators (RRs). Another branch consists of Sho1, which interacts with Cdc42, Ste29 and Ste50 and activates MAPKK Ste11 under osmotic stress conditions. Both branches converge to activate MAPK kinase Pbs2, which leads to phosphorylation of threonine and tyrosine residues on the MAPK HOG1 under stress conditions. Under normal conditions, the phosphorylation of HHKs takes place and the phosphate is relayed to RR via the HPt. The RR activates effector genes and the signaling cascades such as HOG1 MAPK cascade under stress conditions through the dephosphorylation of HHKs (Ssk1), and activation of MAPKK kinases Ssk2 and Ssk22 also takes place. This HOG pathway regulates various stress responses, differentiation and virulence, and represses the mating pathway in *S. cereviseae*. However, the phosphorylation kinetics in *C. neoformans* serotype A is reversed compared to *S. cereviseae*. The phosphorylation of HoG1 takes place by Pbs2 under normal conditions and inhibits the pheromone-MAPK and cAMP-signaling pathways. This phosphorylated form is activated more quickly than the unphosphorylated form under osmotic stress (Bahn *et al*., 2005).

Two genes, *Cl_HP9* and *Cl_BING4CT*, were significantly downregulated in co-cultures. *Cl_HP9* encodes cell cycle protein Blt1 and *Cl_BING4CT* encodes a vegetative incompatibility protein. Vegetative incompatibility is a biological phenomenon occurring in fungal ascomycetes that leads to programmed cell death after the formation of vegetative heterokaryons. These heterokaryons, which are usually non-viable, are formed when vegetative cells from two distinct isolates of the same species fuse. This process is regulated by the heterokaryon incompatibility locus known as *het*. The vegetative incompatibility process that regulates the fusion of vegetative cells is suppressed during mating (Saupe, 2000; Burnett, 2003). Downregulation of *Cl-BING4CT* in co-cultures of two sexually compatible isolates CT-30 and CT-21 of *C. lentis* agrees with a suppression of vegetative incompatibility genes to allow for fusion of sexually compatible mycelia.

Cl_HP9 expression was also significantly downregulated in co-cultures compared to other developmental stages. This gene encodes for cytokinesis protein Blt1p found in the cell cycle during the interphase, contractile ring formation and constriction as well as at division site (Goss *et al*., 2014). It was found that this gene ensures proper timing of the onset and completion of contractile ring constriction as well as septation by recruiting various proteins related to septation

initiation network (SIN) such as Mob1p and Sid2p to the division site. These interact further with various other proteins such as Clp1p and Bgs1p, which in turn regulate cytokinesis and mitotic exit as well as primary septum formation leading to completion of the cell division in fission yeast. In Δ *Blt1p* cells, all these events were delayed including cytokinesis and completion of the cell division cycle leading to cell cycle arrest. In yeast cells, it has been documented that when two cells of opposite mating type are grown together, a cell cycle arrest takes place due to pheromone signaling along with various other physiological changes prior to mating, hence leading to a synchronization of the cells of opposite mating types (Hartwell, 1973). Likewise, in the current study when two sexually compatible isolates of *C. lentis* were grown together, the expression of *Cl_HP9* was repressed, potentially in response to pheromone signaling. It can be speculated that this led to cell cycle arrest in preparation for the cells to fuse. However, the role of this gene needs to be further validated by detecting the presence of the pheromones in the *C. lentis* genome and performing mutational studies for *Cl_HP9.*

Cl_HP3 and *Cl_SART1* also showed significant upregulation in co-cultures suggesting that they may be key components in the signaling pathways during chemotropic interactions. *Cl_HP3* is homologous to the silencing protein Stc1 (siRNA to chromatin). Using RT-qPCR, Bayne *et al*. (2010) showed that Stc1 is required to establish silencing at the *mat2-mat3* loci in fission yeast whereas the maintenance of silencing was independent of Stc1. Silent *mat2-mat3* loci become functional during the mating type switching process in fission yeast (Thon *et al*., 2018). Various silencing mechanisms such as methylation, quelling and repeat-induced point (RIP) mutations have been reported in fungi (Goyon and Faugeron 1989, Romano and Macino 1992, Selker and Stevens 1987). Gene silencing mechanism via RIP mutations are best understood in *N. crassa* (Selker and Stevens 1987).

Cl_SART1 encodes for a SART family protein. The Sart-1 domain-containing protein is a homolog of splicing factor Snu66 in fission yeast. This factor interacts with ubiquitin-like protein Hub1, which plays a role in cell cycle control as well as pre-mRNA splicing. Hub1 mutants show defects in mating in *S. cerevisea* (Dittmarr *et al*., 2002). It is involved in the localisation of cell polarity factors Sph1 and Hbt1. When yeast Δ *hub1* MATa cells were exposed to MATα, there was a disruption of the subcellular localization of both Hbt1 and Sph1, and a defective cell polarization during the formation of mating projections. These mutants were deficient in mating with consistent polarization defects (Dittmarr *et al*., 2002).

Downregulation in perithecial stage 2 was observed for *Cl_Cro1,* which is homologous to *cro1* (actin-cytoskeleton organization) of *P. anserina*. In that species, this gene is expressed during the initial stage of dikaryon formation and then during ascospore maturation. The functional studies of *cro1* in *P. anserina* indicated a role of this gene in septum formation and that this gene is required for the transition from the vegetative to the sexual stage in filamentous fungi. However, the regulation of septum formation during the dikaryotic phase of the sexual stage is not fully understood in ascomycete species. This protein was also found to be homologous to SHE4 (SW15-dependent *HO* expression protein 4) of budding yeasts, which plays a role in mating type switching (Berteaux-Lecellier *et al*., 1998).

In conclusion, several genes that have been associated with the regulation of sexual mating in other ascomycota were identified at QTL *qClMAT3*. These included genes are involved in signalling pathways (*Cl_KH* and *Cl_Hpt*) and in mRNA trafficking (*Cl_HP9*, *Cl_IDC1*, *Cl_SART1*). Their expression patterns agreed with those reported in model species. Some genes, e.g. *Cl_HP9*, *Cl_BING4CT* and *Cl_IDC1* had expression patterns specifically in the mycelial co-cultures, which indicated that they are the most promising candidates associated with mating incompatibility. The expression patterns of these most promising genes need to be confirmed in other compatible isolates belonging to mIG-1 and mIG-2 by performing RT-qPCR for those genes at vegetative and reproductive stages of the compatible isolates. Further validation for their role in mating incompatibility could be achieved using a gene silencing approach or by creating mutants and studying their role in mating through classical mating experiments. It would also be interesting to harvest sufficient stage 1 perithecia to study expression of these genes. The presence of pheromones and their cognate receptors also needs to be determined, which could be done by blasting the conserved domains associated with both α-factor and a-factor pheromones of ascomycete model species against the *C. lentis* genome. This study was mainly focused on 11 genes from a total of 28 genes at QTL *qCLMAT3* using two compatible isolates, so evaluation of expression patterns for the remaining 17 genes at vegetative and reproductive stages is warranted.

6.0 References

Arie T., Kaneko I., Yoshida T., Noguchi M., Nomura Y., and Yamaguchi I. (2000). Mating-type genes from asexual phytopathogenic ascomycetes *Fusarium oxysporium* and *Alternaria alternata*. *Mol Plant Microbe Interact* **13**: 1330–1339.

Armstrong-Cho C., and Banniza S. (2006). *Glomerella truncata* sp. nov., the teleomorph of *Colletotrichum truncatum*. *Mycol Res.* **110**: 951–956.

Armstrong-Cho C., Wang J., Wei Y., and Banniza S. (2012). The infection process of two pathogenic races of *Colletotrichum truncatum* on lentil. *Can J Plant Pathol.* **34**(1): 58–67.

Bahn YS., Kojima K., Cox GM., and Heitman J. (2005). Specialization of the HOG pathway and its impact on differentiation and virulence of Cryptococcus neoformans. *Mol Biol Cell* **16**(5): 2285–2300.

Barsoum E., Martinez P., and Astrom SU. (2010). α3, a transposable element that promotes host sexual reproduction. *Genes Dev.* **24**: 33- 44.

Bayne EH., White SA., Kagansky A., Bijos DA., Sanchez-Pulido L., *et al*. (2010). Stc1: A critical link between RNAi and chromatin modification required for heterochromatin integrity. *Cell* **140**: 666–677.

Berteaux-Lecellier V., Zickler D., Debuchy R., Panvier-Adoutte A., Thompson-Coffe C., and Picard M. (1998). A homologue of the yeast SHE4 gene is essential for the transition between the syncytial and cellular stages during sexual reproduction of the fungus *Podospora anserina*. *EMBO J.***17**(5):1248-58.

Beus EM., Brockenbrough JS., Hong B., and Aris JE. (1994). Yeast NOP2 Encodes an Essential Nucleolar Protein with Homology to a Human Proliferation. *J Cell Biol.* **127**: 1799-1813.

Bhadauria V., MacLachlan R., Pozniak C., Cohen-Skalie A., Li L.*et al*. (2019). Genetic map-guided genome assembly reveals a virulence-governing minichromosome in the lentil anthracnose pathogen *Colletotrichum lentis*. *New Phytol.* **221**(1): 431-445.

Bhadauria V., Ramsay L., Bett KE., and Banniza S. (2017). QTL mapping reveals genetic determinants of fungal disease resistance in the wild lentil species *Lens ervoides*. *Sci Rep.* **7**: 3231-3240.

Bhadauria V., McLachlan R., Pozniak C., and Banniza S. (2015). Candidate effectors contribute to race differentiation and virulence of the lentil anthracnose pathogen *Colletotrichum lentis*. *BMC Genomics* **16**: 628-649.

Buchwaldt L. (2011). Lentil anthracnose. In: Muehlbauer F, Chen W, editors. Compendium of lentil and chickpea diseases. *APS*. 31–34.

Buchwaldt L., and Diederichsen A. (2004). New disease resistant lentil germplasm identified at plant gene resources of Canada (PGRC). *Proceedings of the 5th Canadian Pulse Research Workshop*; London, ON; p. 27.

Buchwaldt L., Dzananovic E., and Durkin J. (2018). Lentil anthracnose: epidemiology, fungicide decision support system, resistance and pathogen races. *Can J Plant Pathol.* **40**(2):189-198.

Buchwaldt L., Shaikh R., Adam J., Tullu A., and Slinkard AE. (2013). Recessive and dominant genes confer resistance to *Colletotrichum truncatum* in cultivated lentil. *Can J Plant Pathol*. **35**:222–231.

Buchwaldt L., Anderson KL., Morall RAA., Gossen BD., and Bernier CC. (2004). Identification of lentil germplasm resistant to *Colletotrichum truncatum* and characterization of two pathogen races. *Phytopathology* **94**: 236- 43.

Buchwaldt L., Gossen B., and Chongo G. (1999). Effect of foliar fungicide application for control of diseases in lentil. *Pest Manage rept*. **100**:277-279.

Buchwaldt L., Morall RAA., Chongo G., and Bernier CC. (1996). Windborne dispersal of *Colletotrichum truncatum* and survival in infested lentil debris. *Phytopathology*. **86**:1193-1198.

Burnett J. 2003. Fungal populations and species. Oxford: Oxford University Press.

Carmell MA., and Hannon GJ. (2004). RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol.* **11**: 214–218.

Chongo G., and Bernier CC. (2000). Effects of host, inoculum concentration, wetness duration, growth stage, and temperature on anthracnose of lentil. *Plant Dis.* **84**: 544-548.

Chongo G., Gossen BD., and Bernier CC. (2002). Infection by *Colletotrichum truncatum* in resistant and susceptible lentil genotypes. *Can J Plant Pathol.* **24**: 81–85.

Cisar CR., and TeBeest DO. (1999). Mating system of the filamentous ascomycete, *Glomerella cingulata*. *Curr Genet.* **35**: 127- 133.

Coppin E., Debuchy R., Arnaise S., and Picard M. (1997). Mating types and sexual development in filamentous ascomycetes. *Mol Biol Rev.* **61**: 411–428.

Coppin E., Arnaise S., Contamine V., and Picard M. (1993). Deletion of mating type sequences in *Podospora anserine* abolishes mating without affecting vegetative functions and sexual differentiation. *Mol Gen Genet*. **241**: 409-414.

Crouch J., O'Connell R., Gan P., Buiate E., Torres MF., *et al*. (2014). The Genomics of *Colletotrichum*. In: Dean R., Lichens-Park A., Kole C. (eds*) Genomics of Plant-Associated Fungi: Monocot Pathogens*. *Springer, Berlin, Heidelberg*.

Damm U., Sato T., Alizadeh A., Groenewald JZ., and Crous PW. (2019). The *Colletotrichum dracaenophilum*, *C. magnum* and *C. orchidearum* species complexes. *Stud Mycol.* **92**: 1–46.

Damm U., O'Connell RJ., Groenewald JZ., and Crous PW. (2014). The *Colletotrichum destructivum* species complex – hemibiotrophic pathogens of forage and field crops. *Stud Mycol.* **79**: 49–84.

Dean R., Van Kan JAL., Pretorius ZA., Hammond-Kosack KE., Di Pietro A., *et al*. (2012). The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol.* **13**(4): 414–430.

Debuchy R., Berteaux-Lecellier V., and Silar P. (2010). Mating systems and sexual morphogenesis in ascomycetes. In: Borkovich KA., Ebbole D.J (eds) *Cellular and molecular biology of filamentous fungi*. ASM Press, Washington, DC, pp: 501–535.

Dittmar GA., Wilkinson CR., Jedrzejewski PT., and Finley D. (2002). Role of a ubiquitin-like modification in polarized morphogenesis. *Science* **295**: 2442–2446.

Doughan B., and Rollins JA. (2016). Characterization of *MAT* Gene Functions in the Life Cycle of *Sclerotinia sclerotiorum* reveals a Lineage-Specific *MAT* Gene Functioning in Apothecium Morphogenesis. *Fungal Biol*. **120**: 1105–1117.

Du M., Schardl CL., Nuckles EM., and Vaillancourt LJ. (2005). Using mating-type gene sequences for improved phylogenetic resolution of *Colletotrichum* species complexes. *Mycologia* **97**(3): 641–658.

Duong AT., Debeer WZ., Wingfield DB., and Wingfield JM. (2013). Characterization of the mating–type genes in *Leptographium procerum* and *Leptographium profanum*. *Fungal Biol.* **117**: 411-421.

Durkin J., Bisset J., Pahlavani M., Mooney B., and Buchwaldt L. (2015). IGS minisatellites useful for race differentiation in *Colletotrichum lentis* and a likely site of small RNA synthesis affecting pathogenicity. *PLoS One* **10**(9): e 0137398.

Dyer P., Inderbitzin P., and Debuchy R. (2016.). Mating-type structure, function, regulation and evolution in the Pezizomycotina. In: Wendland J., editor. Growth, Differentiation and Sexuality. *The Mycota* **1**: 351–385.

Espagne E., and Barreau C. (2002). HET -E and HET-D belong to a new subfamily of WD40 proteins involved in vegetative incompatibility specificity in the fungus *Podospora anserina. Genetics* **161**: 71–81.

Exil VJ., Silva Avila D., Benedetto A., Exil EA., Adams MR., *et al*. (2010). Stressed-induced TMEM135 protein is part of a conserved genetic network involved in fat storage and longevity regulation in *Caenorhabditis elegans*. *PLoS One* **5**: 14228.

Feldbrugge M., Kamper J., Steinberg G., and Kahmann R. (2004). Regulation of mating and pathogenic development in *Ustilago maydis*. *Curr. Opin. Microbiol*. **7**: 666–672.

García-Serrano M., Laguna EA., Simpson J., and Rodríguez-Guerra R. (2008). Analysis of the *MAT1-2-1* gene of *Colletotrichum lindemuthianum*. *Mycoscience* **49**(5):312–317.

Gautier V., Tong L., and Nguyen TS. (2018). PaPro1 and IDC4, two genes controlling stationary phase, sexual development and cell degeneration in *Podospora anserina*. *J. Fungi* **4**:85.

Gela TS., Banniza S., and Vandenberg A. (2020). Lack of effective resistance to the virulent race of *Colletotrichum lentis* in *Lens culinaris* Medikus subsp. culinaris. *Plant Genet Res.* **18**(2):81-87.

Gelin-Licht R., Paliwal S., Conlon P., Levchenko A., and Gerst JE. (2012). Scp160-dependent mRNA trafficking mediates pheromone gradient sensing and chemotropism in yeast. *Cell Rep*. **1**(5):483-94.

Glass NL., Vollmer SJ., Staben C., Metzenberg RL., and Yanofsky C. (1988). DNA of the two mating-type alleles of *Neurospora crassa* are highly dissimilar. *Science* **241**: 570–573.

Goss JW., Kim S., Bledsoe H., and Pollard TD. (2014). Characterization of the roles of Blt1p in fission yeast cytokinesis. *Mol Biol Cell* **25**(13):1946-57.

Gossen BD., Anderson K., and Buchwaldt L. (2009). Host specificity of *Colletotrichum truncatum* from lentil. *Can J Plant Pathol.* **31**: 65-73.

Goyon C., and Faugeron G. (1989). Targeted transformation of *Ascobolus immerses* and de novo methylation of the resulting duplicated DNA sequences. *Mol Cell Biol*. **9**(7): 2818–2827.

Guo M., Aston C., Burchett S., Dyke C., Fileds S., *et al*. (2003). The yeast G protein α subunit Gpa1 transmits a signal through an RNA binding effector protein Scp160. *Mol Cell* **12**, 517–524.

Haber JE. (2012). Mating type genes and MAT switching in *Saccharomyces cerevisiae*. *Genetics* **191**: 33-64.

Hartwell LH. (1973). Synchronization of haploid yeast cell cycles, a prelude to conjugation. *Exp. Cell Res.* **76**:111-117.

Hiruma K., Gerlach N., Sacristán S., Nakano RT., Hacquard S., *et al*. (2016). Root endophyte *Colletotrichum tofieldiae* confers plant fitness benefits that are phosphate status dependent. *Cell* **165**(2): 464 -474.

Huang F., Chen GQ., Hou X., Fu YS., Cai L., *et al*. (2013). *Colletotrichum* species associated with cultivated citrus in China. *Fungal Divers.* **61**: 61–74.

Jamet-Vierny C., Debuchy R., Prigent M., and Silar P. (2007). IDC1, a pezizomycotina-specific gene that belongs to the PaMpk1 MAP kinase transduction cascade of the filamentous fungus *Podospora anserina*. *Fungal Genet Biol.* **44**(12):1219-30.

Jonkers W., Leeder AC., Ansong C., Wang Y., Yang F., *et al*. (2014). HAM-5 functions as a MAP Kinase scaffold during cell fusion in Neurospora crassa. *PLoS Genet.* **10**(11):1004783.

Khazaei H., Podder R., Caron C., Kundu S., Diapari M., Vanderberg A., and Bett KE. (2017). Marker-trait association analysis of iron and zinc concentration in lentil (*Lens culinaris* Medik.) seeds. *Plant Genome* **10**(2).

Kim H., and Borkovich KA. (2006). Pheromones are essential for male fertility and sufficient to direct chemotropic polarized growth of trichogynes during mating in *Neurospora crassa*. *Eukaryot. Cell* **5**: 544 -554.

Kim H., Wright SJ., Park G., Ouyang S., Krystofova S., and Borkovich KA. (2012). Roles for receptors, pheromones, G proteins and mating type genes during sexual reproduction in *Neurospora crassa*. *Genetics.***190**:1389–1404.

Kumar G., and Baojun Xu. (2017). Polyphenol-rich lentils and their health promoting effects. *Int J Mol Sci*. **18**(11):2390.

Lee J., Leslie JF., and Bowden RL. (2008). Expression and function of sex pheromones and receptors in the homothallic ascomycete *Gibberella zeae*. *Eukaryot Cell* **7**:1211–1221.

Livak KJ., and Schmittgen TD. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods* **25**(4): 402-408.

Liu F., Cai L., Crous PW., and Damm U. (2014). The *Colletotrichum gigasporum* species complex. *Persoonia - Mol Phylogeny Evol Fungi* **33**: 83–97.

Maksimov V., Wäneskog M., Rodriguez A., and Bjerling P. (2017). Stress sensitivity of a fission yeast strain lacking histidine kinases is rescued by the ectopic expression of Chk1 from *Candida albicans. Curr Genet* **63**:343–357.

Marat AL., Dokainish H., and McPherson PS. (2011). DENN domain proteins: Regulators of Rab GTPases. *J. Biol. Chem.* **286:**13791–13800.

Marin-Felix Y., Groenewald JZ., Cai L., Chen Q., Marincowitz S., *et al*. (2017). Genera of phytopathogenic fungi: GOPHY 1. *Stud Mycol*. **86**: 99–216.

Mayrhofer S., Weber JM., and Poggeler S. (2006). Pheromones and pheromone receptors are required for proper sexual development in the homothallic ascomycete *Sordaria macrospora*. *Genetics*. **172**: 1521-1533.

Menat J., Armstrong-Cho C., and Banniza S. (2016). Lack of evidence for sexual reproduction in field populations of *Colletotrichum lentis*. *Fungal Ecol.* **20**: 66–74.

Menat J., Cabral AL., Vijayan P., Wei Y., and Banniza S. (2012). *Glomerella truncata*: another *Glomerella* species with an atypical mating system. *Mycologia* **104**: 641-649.

Mohanan VC., Chandarana PM., Chattoo BB., Patkar RN., and Manjrekar J. (2017). Fungal histidine phosphotransferase plays a crucial role in photomorphogenesis and pathogenesis in *Magnaporthe oryzae*. *Front Chem*. **5**: 31.

Moller M., and Stukenbrock EH. (2017). Evolution and genome architecture in fungal plant pathogens. *Nat Rev Microbiol.* **15**(12): 771.

Morall RAA. (1997). Evolution of lentil diseases over 25 years in western Canada. *Can J Plant Pathol*. **19**: 197-207.

Morrall RAA. (1988). A new disease of lentil induced by *Colletotrichum truncatum* in Manitoba. *Plant Dis*. **72**: 994.

Morrall RAA., and Pedersen EA. (1991). Discovery of lentil anthracnose in Saskatchewan in 1990. *Can Plant Dis Surv.* 71: 105–106.

Moseley BJ., Mayeux A., Paoletti A., and Nurse P. (2009). A spatial gradient coordinates cell size and mitotic entry in fission yeast. *Nature* **459**:857–860.

Nakayama N., Miyajima A., and Arai K. (1985). Nucleotide sequences of STE2 and STE3, cell type-specific sterile genes from *Saccharomyces cerevisiae*. *The EMBO journal* **4**: 2643.

Nelson WJ. (2003). Adaptation of core mechanisms to generate cell polarity. *Nature* **422**(6933): 766-774.

O'Connell RJ., and Panstruga R. (2006). Tete a tete inside a plant cell: establishing compatibility between plants and biotrophic fungi and oomycetes. *New Phytol.* **171**(4): 699–718.

O'Connell R., Perfect S., Hughes B., Carzaniga R., Bailey J., and Green J. (2000). Dissecting the cell biology of *Colletotrichum* infection processes. In: *Colletotrichum*: *host specificity, pathology, and host-pathogen interaction*. APS Press, St. Paul, Minnesota, USA, pp: 57-77.

Ramírez-Peinado S., Ignashkova TI., Van Raam BJ., Baumann J., Sennott EL., *et al*. (2017). TRAPPC13 modulates autophagy and the response to Golgi stress. *J Cell Sci*. **130**: 2251–2265.

Reblova M., Gams W., and Seifert KA. (2011). Monilochaetes and allied genera of the Glomerellales, and a reconsideration of families in the Microascales. *Stud Mycol.* **68**(1):163–191.

Robinson JA., and Natvig OD. (2019). Diverse members of the *Xylariales* lack canonical mating type regions. *Fungal Genet Biol.* **122**: 47-52.

Rodenburg SY., Terhem RB., Veloso J., Stassen JH., and Van Kan JA. (2018). Functional analysis of mating type genes and transcriptome analysis during fruiting body development of *Botrytis cinerea*. *mBio*. **9**: 01939-17.

Rodriguez-Guerra R., Ramírez-Rueda MT., Cabral-Enciso M., García-Serrano, M., Lira-Maldonado Z., *et al*. (2005). Heterothallic mating observed between Mexican isolates of *Glomerella lindemuthiana*. *Mycologia* **97**(4):793–803.

Romano N., and Macino G. (1992). Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol microbiol*. **6**(22):3343-53.

Rotin D., and Kumar S. (2009). Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol*. **10**: 398-409.

Salamov AA., and Solovyev VV. (2000). Ab initio gene finding in drosophila genomic DNA. *Genome Res.* **10:**516–522.

Saupe SJ. (2000). Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. *Microbiol Mol Biol Rev.* **64**(3):489.

Seibel C., Tisch D., Kubicek CP., and Schmoll M. (2012). ENVOY is a major determinant in regulation of sexual development in *Hypocrea jecorina* (*Trichoderma reesei*). *Eukaryot Cell* **11**: 885–895.

Selker L., and Stevens JN. (1987). Signal for DNA methylation associated with tandem duplication in Neurospora crassa. *Mol Cell Biol*. **7**(3):1032-1038.

Shaikh R., Diederichsen A., Harrington M., Adam J., Conner RL., and Buchwaldt L. (2013). New sources of resistance to *Colletotrichum truncatum* race Ct0 and Ct1 in *Lens culinaris* Medikus subsp. *culinaris* obtained by single plant selection in germplasm accessions. *Genetic Resour Crop Evol.* **60**: 193–201.

Song K, Mach K E, Chen CY., Reynolds T., and Albright CF. (1996). CFA novel suppressor of *ras1* in fission yeast, *byr4*, is a dosage-dependent inhibitor of cytokinesis. *J Cell Biol.* **133**: 1307-131.

Sutton BC. (1992). The genus *Glomerella* and its anamorph *Colletotrichum*. In: *Colletotrichum*: biology, pathology and control (eds. J.A. Bailey and M.J. Jeger). *CAB International, Wallingford*: 1-26.

Taylor JW. (1995). Molecular phylogenetic classification of fungi. *Arch. Med. Res*. **26**: 307-314.

Taylor P., Lindbeck K., Chen W., and Ford R. (2007). Lentil diseases. In: Yadav SS, McNeil D, Stevenson PC (eds) *Lentil: an ancient crop for modern times. Springer, Dordrecht*.

Thon G., Maki T., Haber JE., and Iwasaki H. (2019). Mating-type switching by homology-directed recombinational repair: A matter of choice. *Curr genet.* **65**(2):351-62.

Tullu A., Bett KE., Banniza S., Vail S., and Vandenberg A. (2013). Widening the genetic base of cultivated lentil through hybridization of *Lens culinaris* 'Eston'and *L. Ervoides* accession IG72815. *Can J Plant Sci* **93**: 1037–1047.

Tullu A., Buchwaldt L., Lulsdorf M., Banniza S., Barlow B. *et al.* (2006). Sources of resistance to anthracnose (*Colletotrichum truncatum*) in wild *Lens* species. *Genet Resour Crop Evol.* **53**: 111–119.

Tullu A., Buchwaldt L., Warkentin T., Taran B., and Vandenberg A. (2003). Genetics of resistance to anthracnose and identification of AFLP and RAPD markers linked to the resistance gene in PI 320937 germplasm of lentil (*Lens culinaris* Medikus). *Theor Appl Genet*. **106**:428–434.

Turgeon BG., and Yoder OC. (2000). Proposed nomenclature for mating type genes of filamentous ascomycetes. *Fungal Genet Biol.* **31**:1–5.

Untergasser A., Cutcutache I., Koressaar T., Ye J., *et al*. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **40**(15): e115.

Vaillancourt, Wang J., and Hanau R. (2000). Genetic regulation of sexual compatibility in *Glomerella graminicola.* In: Prusky D, Freeman S, Dickman M, eds. *Colletotrichum*: *host specificity, pathology and host-pathogen interaction*, p 23–43.

Vanderberg A., Kiehen FA., Vera C., Gaudiel R., Buchwaldt L., *et al*. (2002). 'CDC Robin' lentil. *Can J Plant Sci.* **82**: 111-112.

Warale R. (2015). Determining genetic diversity and regulation of sexual compatibility in *Colletotrichum lentis* Damm, the causal agent of anthracnose on *Lens culinaris* (Medik.). MSc. Thesis, Department of Plant Science, University of Saskatchewan.

Wheeler HE. (1956). Linkage groups in Glomerella. *Am J Bot.* **43**:1-6.

Wilken MP., Steenkamp TE., Nest AM., Wingfield JM., Beer WZ., and Wingfield DB. (2018). Unexpected placement of the MAT 1-1-2 gene in the MAT1-2 idiomorph of *Thielavopsis*. *Fungal Gen and Bio.* **113**: 32-41.

Wilkinson CR., Dittmar GA., Ohi MD., Uetz P., Jones N. and Finley D. (2004). Ubiquitin-like protein Hub1 is required for pre-mRNA splicing and localization of an essential splicing factor in fission yeast. *Curr Biol* **14**: 2283—2288.

Xu S., Christensen MJ., and Li Y. (2017). Pathogenicity and characterization of *Colletotrichum lentis* A causal agent of anthracnose in common vetch (Vicia sativa). *Eur J Plant Pathol.* **149**(3): 719-731.

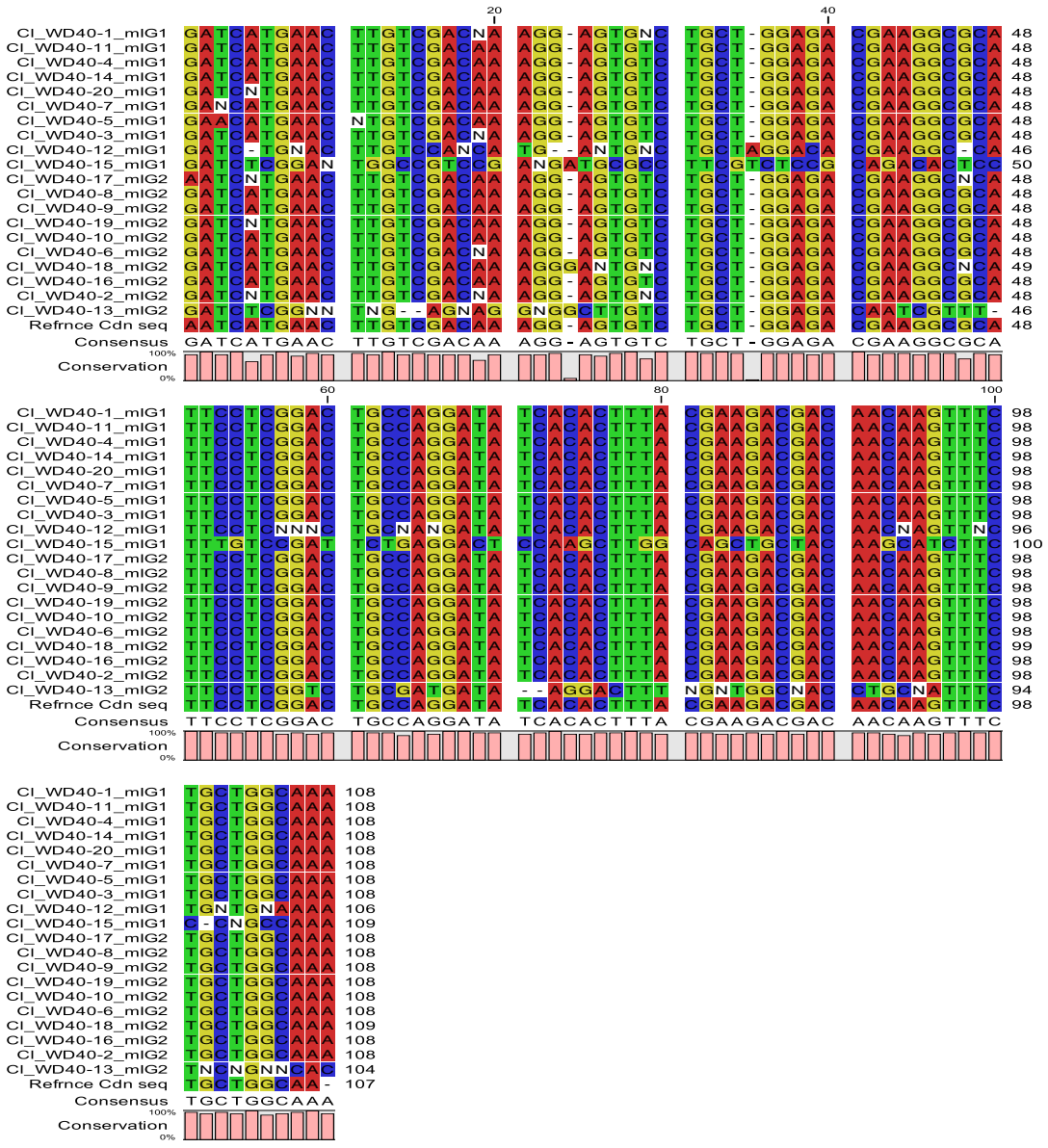
Xue C., Hsueh YP., Heitman J. (2008). Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. *FEMS Microbiol Rev.* **32**(6):1010-32.

Yu Y., Amich J., Will C., Eagle CE., Dyer PS., and Krappmann S. (2017). The novel *Aspergillus fumigatus MAT1-2-4* mating-type gene is required for mating and cleistothecia formation. *Fungal Genet Biol*. **108**: 1–12.

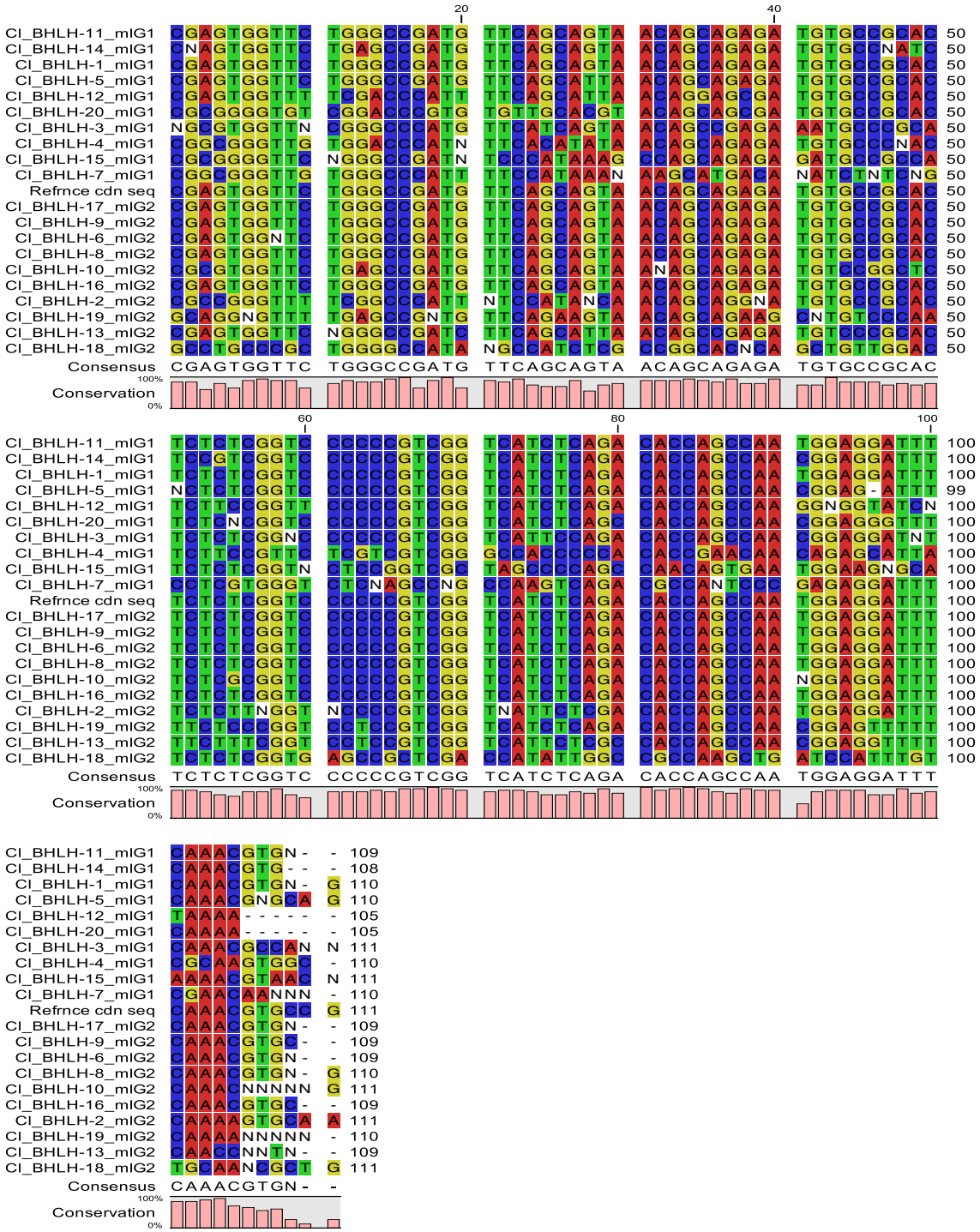
Yun SH., Berbee ML., Yoder C., and Turgeon BG. (1999). Evolution of the fungal self-fertile reproductive lifestyle from self-sterile ancestors. *Proc Natl Acad Sci.* **96**: 5592–5597.

7.0 Appendix

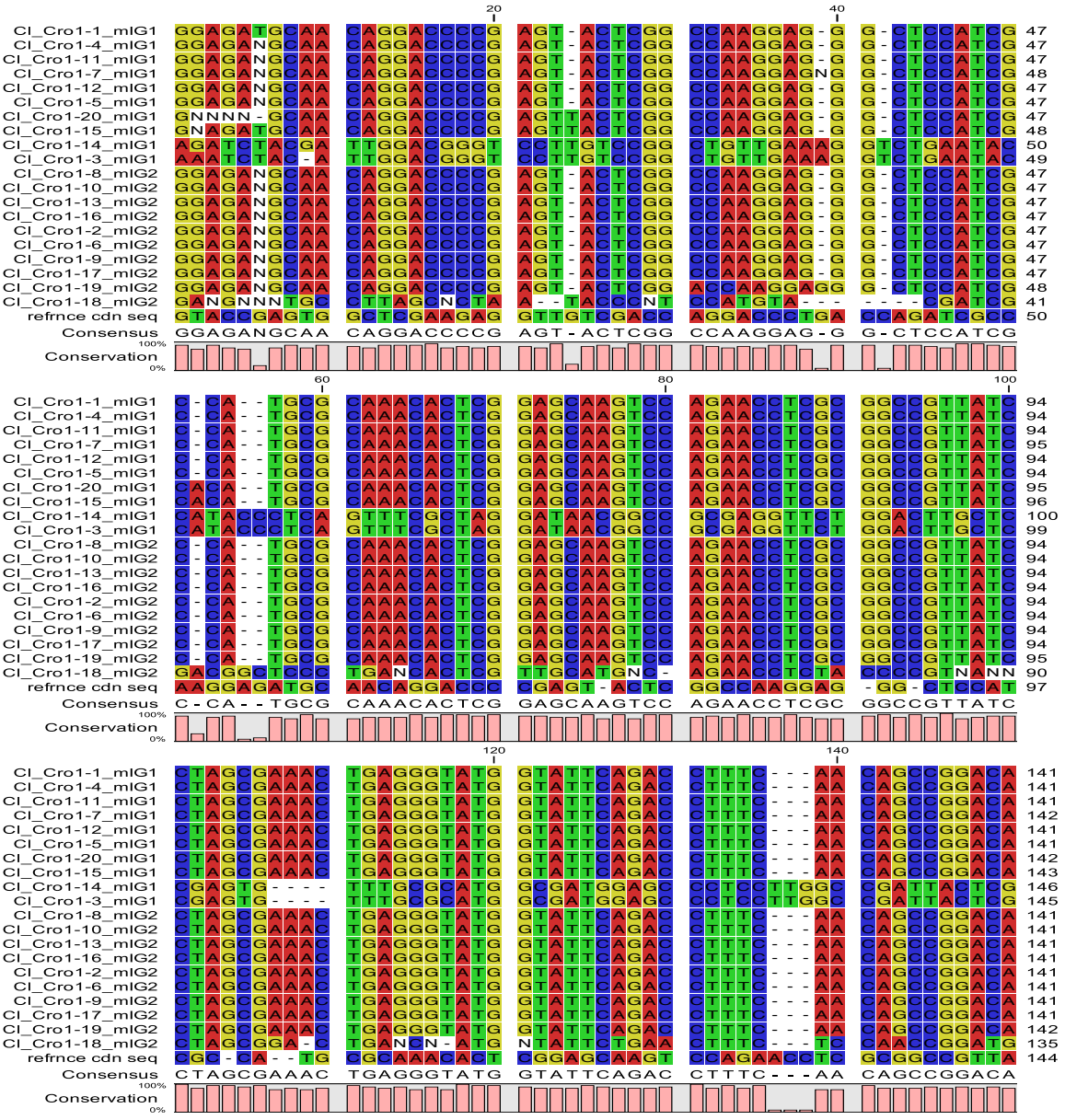
Appendix 1: Sequence alignment of *Cl_WD40* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.

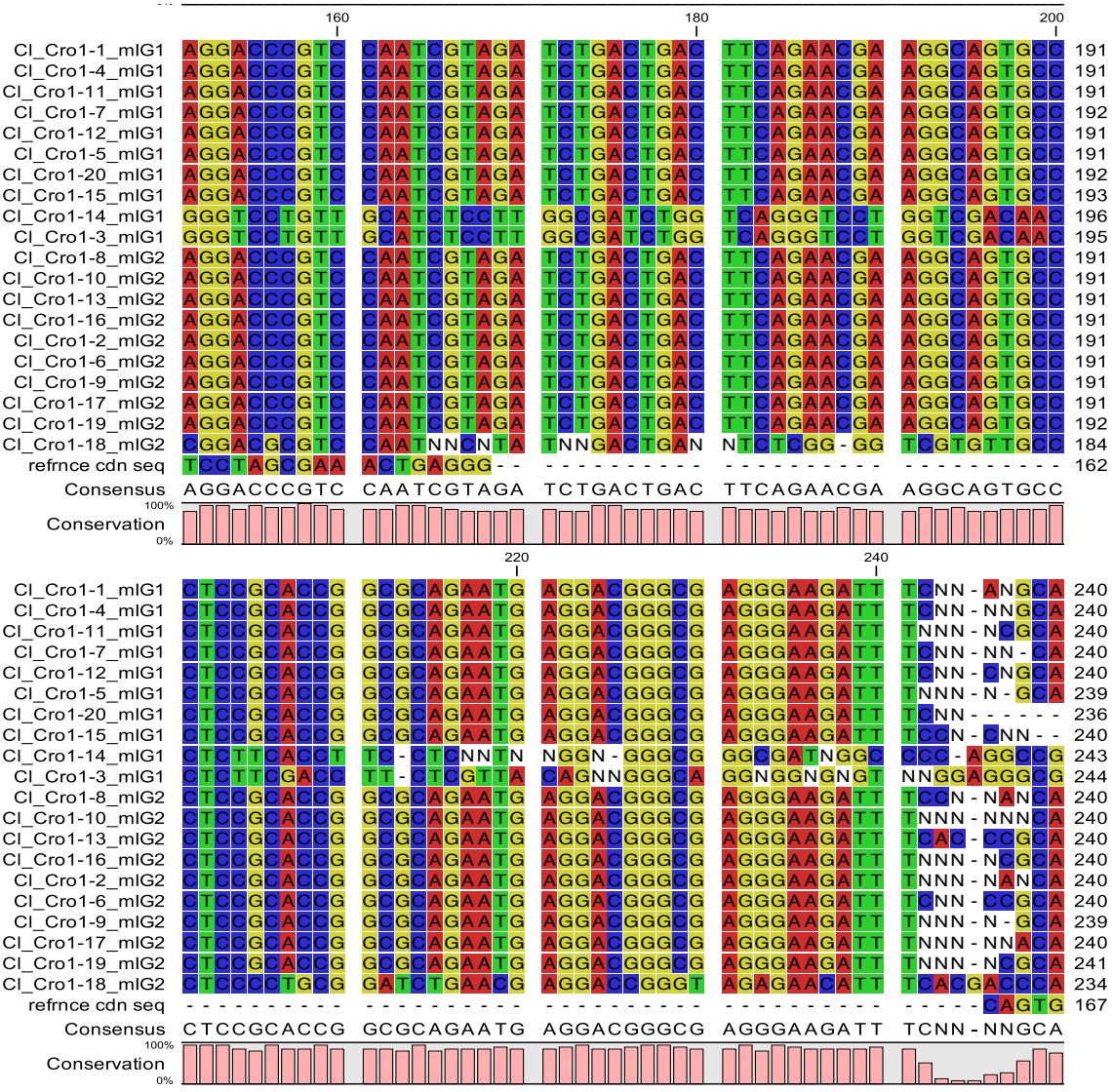


Appendix 2: Sequence alignment of *Cl_BHLH* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.

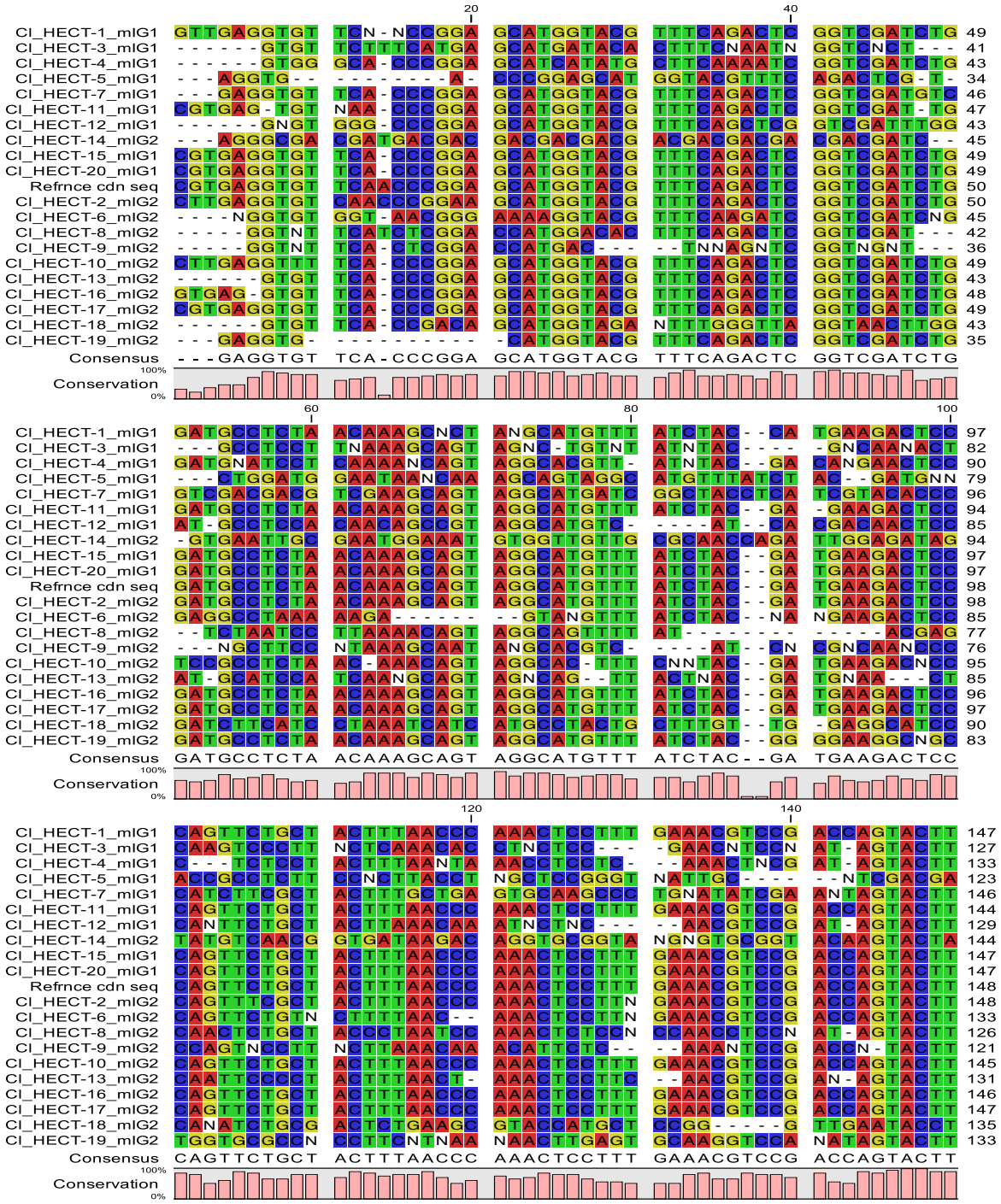


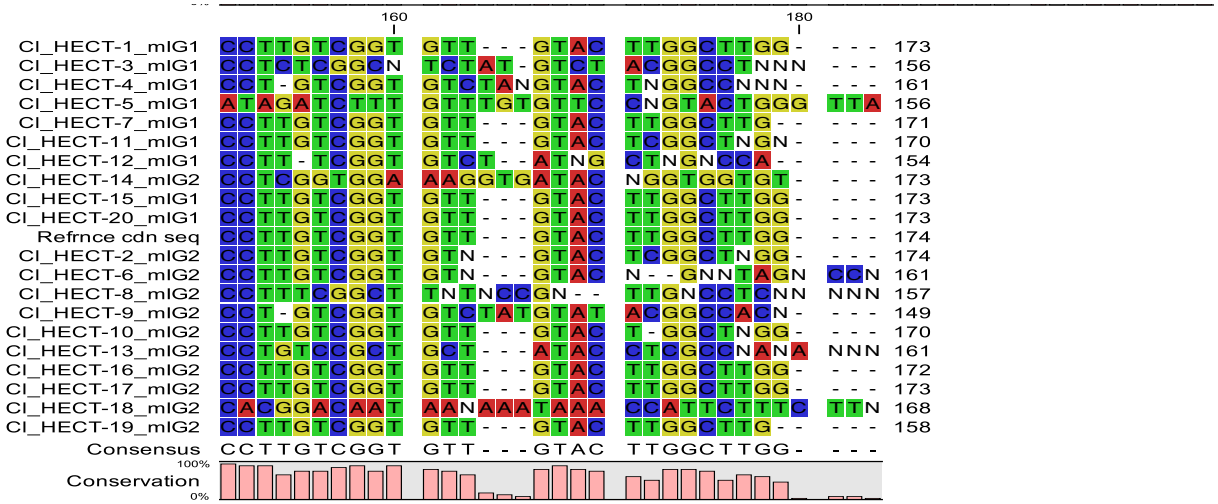
Appendix 3: Sequence alignment of *Cl_Cro1* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software



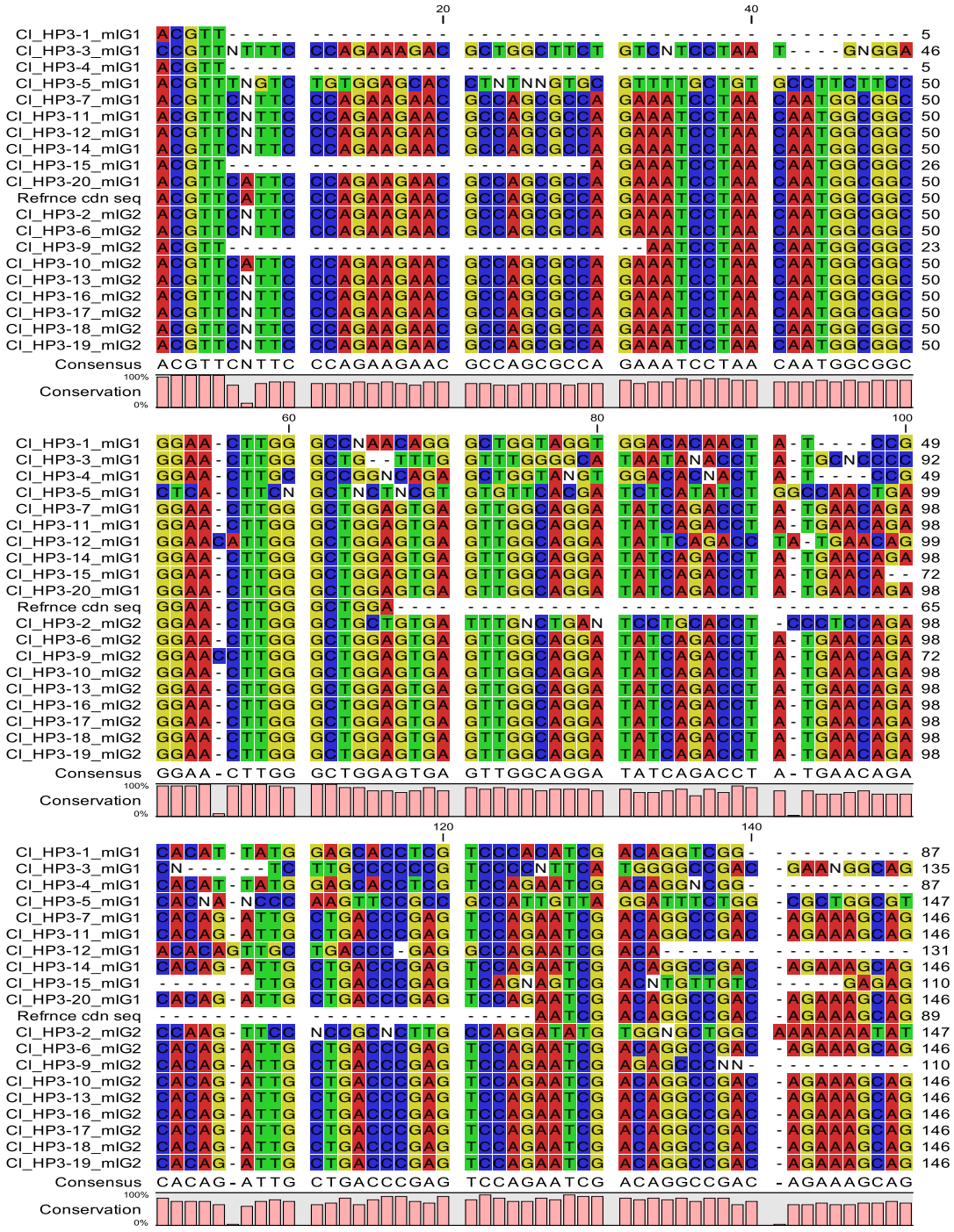

160
180
200
Cl_Cro1-1_mIG1 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-4_mIG1 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-11_mIG1 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-7_mIG1 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 192
Cl_Cro1-12_mIG1 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-5_mIG1 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-20_mIG1 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 192
Cl_Cro1-15_mIG1 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 193
Cl_Cro1-14_mIG1 GGGTCCTGTT GCATCTCCTT GGCGATCTGG TCAGGGTCCT GGTCGACAAC 196
Cl_Cro1-3_mIG1 GGGTCCTGTT GCATCTCCTT GGCGATCTGG TCAGGGTCCT GGTCGACAAC 195
Cl_Cro1-8_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-10_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-13_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-16_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-2_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-6_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-9_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-17_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 191
Cl_Cro1-19_mIG2 AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC 192
Cl_Cro1-18_mIG2 CGGACGCGTC CAATNNCNTA TNNGACTGAN NTCTCGG-GG TCGTGTTGCC 184
refrnce cdn seq TCCTAGCGAA ACTGAGGG-- ---------- ---------- ---------- 162
Consensus AGGACCCGTC CAATCGTAGA TCTGACTGAC TTCAGAACGA AGGCAGTGCC
100%
Conservation
0%
220
240
Cl_Cro1-1_mIG1 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCNN-ANGCA 240
Cl_Cro1-4_mIG1 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCNN-NNGCA 240
Cl_Cro1-11_mIG1 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TNNN-NCGCA 240
Cl_Cro1-7_mIG1 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCNN-NN-CA 240
Cl_Cro1-12_mIG1 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCNN-CNGCA 240
Cl_Cro1-5_mIG1 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TNNN-N-GCA 239
Cl_Cro1-20_mIG1 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCNN------ 236
Cl_Cro1-15_mIG1 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCCN-CNN-- 240
Cl_Cro1-14_mIG1 CTCTTCACCT TC-CTCNNTN NGGN-GGGCG GGCGATNGGC CCC-AGGCCG 243
Cl_Cro1-3_mIG1 CTCTTCGACC TT-CTCGTTA CAGNNGGGCA GGNGGNGNGT NNGGAGGGCG 244
Cl_Cro1-8_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCCN-NANCA 240
Cl_Cro1-10_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TNNN-NNNCA 240
Cl_Cro1-13_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCAC-CCGCA 240
Cl_Cro1-16_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TNNN-NCGCA 240
Cl_Cro1-2_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TNNN-NANCA 240
Cl_Cro1-6_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCNN-CCGCA 240
Cl_Cro1-9_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TNNN-N-GCA 239
Cl_Cro1-17_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TNNN-NNACA 240
Cl_Cro1-19_mIG2 CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TNNN-NCGCA 241
Cl_Cro1-18_mIG2 CTCCCCTGCG GATCTGAACG AGGACCGGGT AGAGAACATT TCACGACCCA 234
refrnce cdn seq ---------- ---------- ---------- ---------- -----CAGTG 167
Consensus CTCCGCACCG GCGCAGAATG AGGACGGGCG AGGGAAGATT TCNN-NNGCA
100%
Conservation
0%

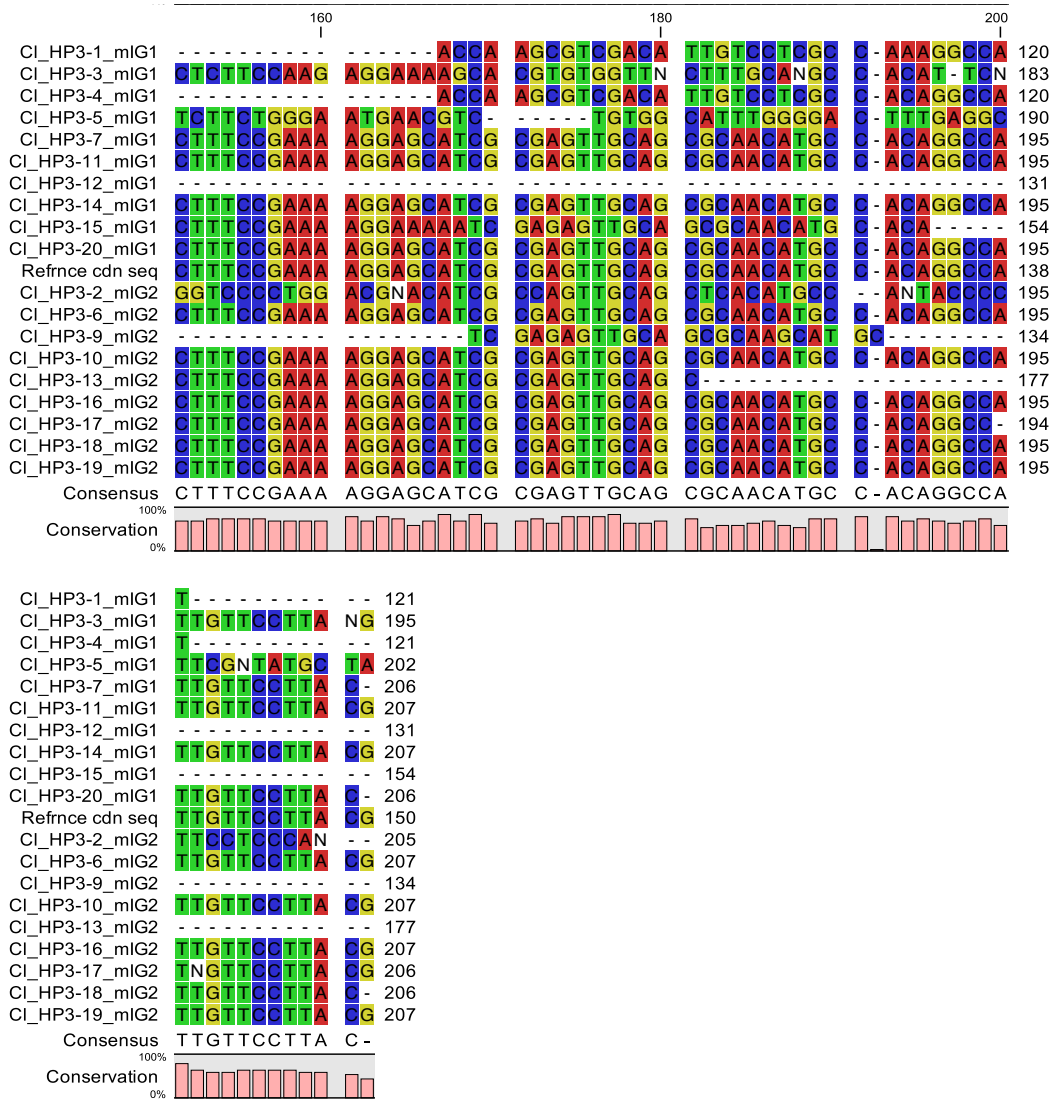
Appendix 4: Sequence alignment of *Cl_HECT* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.




160
180
Cl_HECT-1_mIG1 CCTTGTCGGT GTT---GTAC TTGGCTTGG- ---173
Cl_HECT-3_mIG1 CCTCTCGGCN TCTATGTCT ACGGCCTNNN ---156
Cl_HECT-4_mIG1 CCT-GTCGGT GTCTANGTAC TNGGCCNNN- ---161
Cl_HECT-5_mIG1 ATAGATCTTT GTTTGTGTTC CNGTACTGGG TTA156
Cl_HECT-7_mIG1 CCTTGTCGGT GTT---GTAC TTGGCTTG-- ---171
Cl_HECT-11_mIG1 CCTTGTCGGT GTT---GTAC TCGGCTNGN- ---170
Cl_HECT-12_mIG1 CCTT-TCGGT GTCT--ATNG CTNGNCCA-- ---154
Cl_HECT-14_mIG2 CCTCGGTGGA AAGGTGATAC NGGTGGTGT- ---173
Cl_HECT-15_mIG1 CCTTGTCGGT GTT---GTAC TTGGCTTGG- ---173
Cl_HECT-20_mIG1 CCTTGTCGGT GTT---GTAC TTGGCTTGG- ---173
Refrnce cdn seq CCTTGTCGGT GTT---GTAC TTGGCTTGG- ---174
Cl_HECT-2_mIG2 CCTTGTCGGT GTN---GTAC TCGGCTNGG- ---174
Cl_HECT-6_mIG2 CCTTGTCGGT GTN---GTAC N--GNNTAGN CCN161
Cl_HECT-8_mIG2 CCTTTCGGCT TNTNCCGN-- TTGNCCTCNN NNN157
Cl_HECT-9_mIG2 CCT-GTCGGT GTCTATGTAT ACGGCCACN- ---149
Cl_HECT-10_mIG2 CCTTGTCGGT GTT---GTAC T-GGCTNGG- ---170
Cl_HECT-13_mIG2 CCTGTCCGCT GCT---ATAC CTCGCCNANA NNN161
Cl_HECT-16_mIG2 CCTTGTCGGT GTT---GTAC TTGGCTTGG- ---172
Cl_HECT-17_mIG2 CCTTGTCGGT GTT---GTAC TTGGCTTGG- ---173
Cl_HECT-18_mIG2 CAGGACAAT AANAAATAAA CCATTCTTTC TTN168
Cl_HECT-19_mIG2 CCTTGTCGGT GTT---GTAC TTGGCTTG-- ---158
Consensus CCTTGTCGGT GTT---GTAC TTGGCTTGG- ---
100%
Conservation
0%

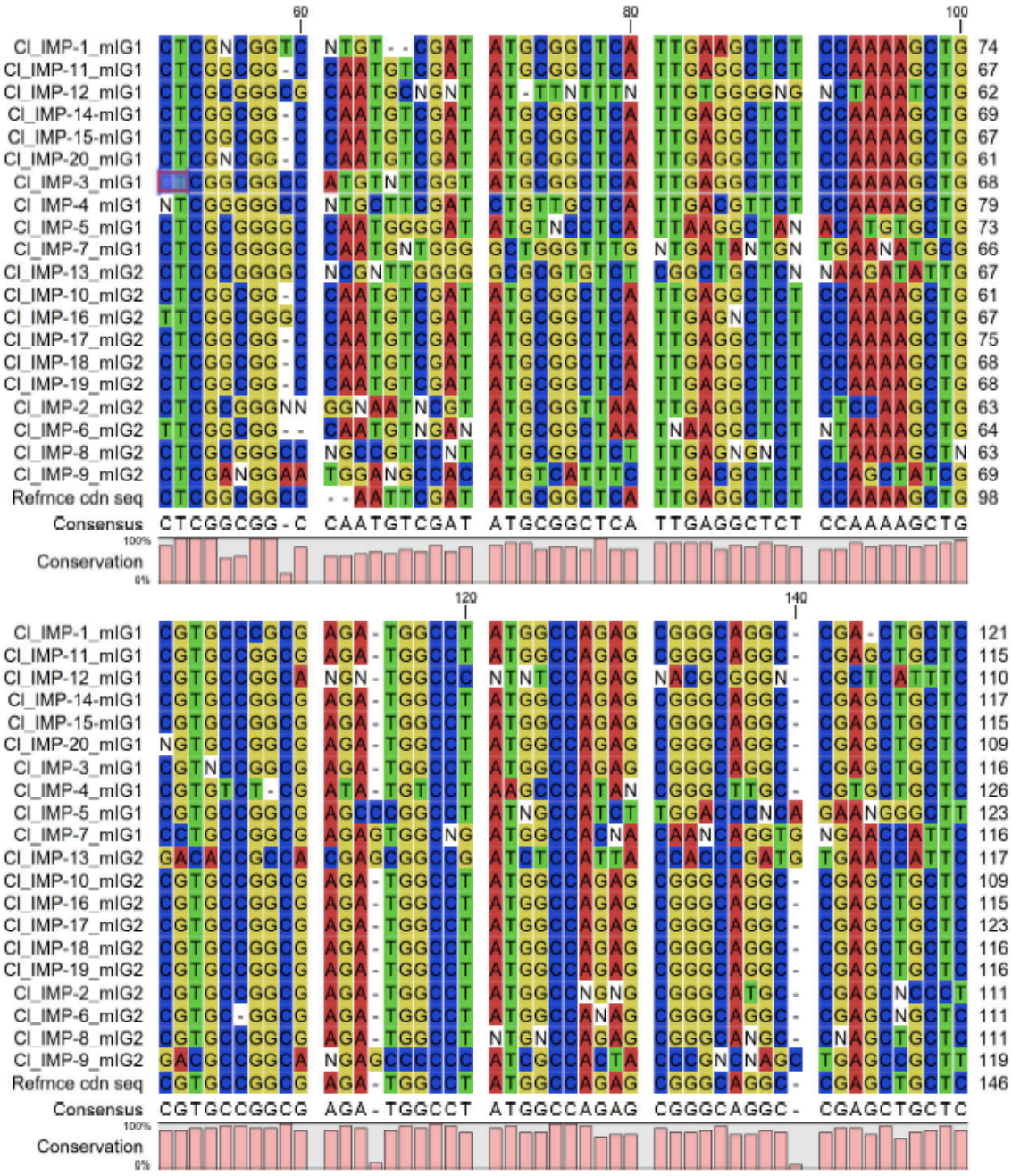
Appendix 5: Sequence alignment of *Cl_HP3* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.

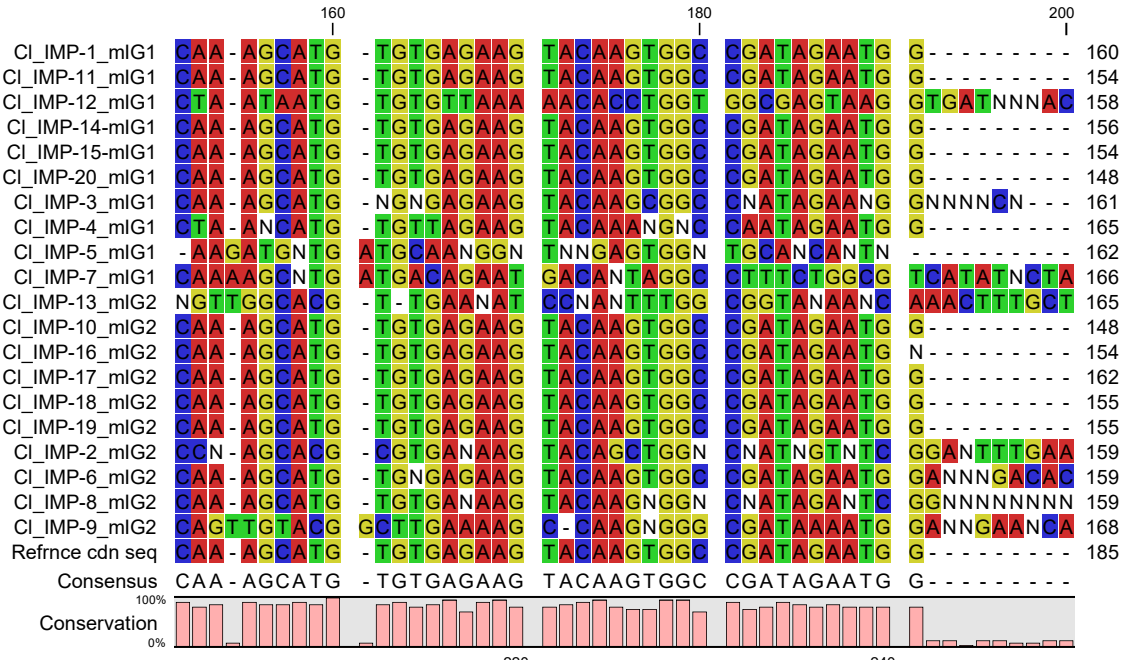



160
180
200
Cl_HP3-1_mIG1 120
Cl_HP3-3_mIG1 183
Cl_HP3-4_mIG1 120
Cl_HP3-5_mIG1 190
Cl_HP3-7_mIG1 195
Cl_HP3-11_mIG1 195
Cl_HP3-12_mIG1 131
Cl_HP3-14_mIG1 195
Cl_HP3-15_mIG1 154
Cl_HP3-20_mIG1 195
Refrnce cdn seq 138
Cl_HP3-2_mIG2 195
Cl_HP3-6_mIG2 195
Cl_HP3-9_mIG2 134
Cl_HP3-10_mIG2 195
Cl_HP3-13_mIG2 177
Cl_HP3-16_mIG2 195
Cl_HP3-17_mIG2 194
Cl_HP3-18_mIG2 195
Cl_HP3-19_mIG2 195
Consensus CTTTCCGAAA AGGAGCATCG CGAGTTGCAG CGCAACATGC C-ACAGGCCA
100%
Conservation
0%
Cl_HP3-1_mIG1 121
Cl_HP3-3_mIG1 195
Cl_HP3-4_mIG1 121
Cl_HP3-5_mIG1 202
Cl_HP3-7_mIG1 206
Cl_HP3-11_mIG1 207
Cl_HP3-12_mIG1 131
Cl_HP3-14_mIG1 207
Cl_HP3-15_mIG1 154
Cl_HP3-20_mIG1 206
Refrnce cdn seq 150
Cl_HP3-2_mIG2 205
Cl_HP3-6_mIG2 207
Cl_HP3-9_mIG2 134
Cl_HP3-10_mIG2 207
Cl_HP3-13_mIG2 177
Cl_HP3-16_mIG2 207
Cl_HP3-17_mIG2 206
Cl_HP3-18_mIG2 206
Cl_HP3-19_mIG2 207
Consensus TTGTTCCTTA C-
100%
Conservation
0%

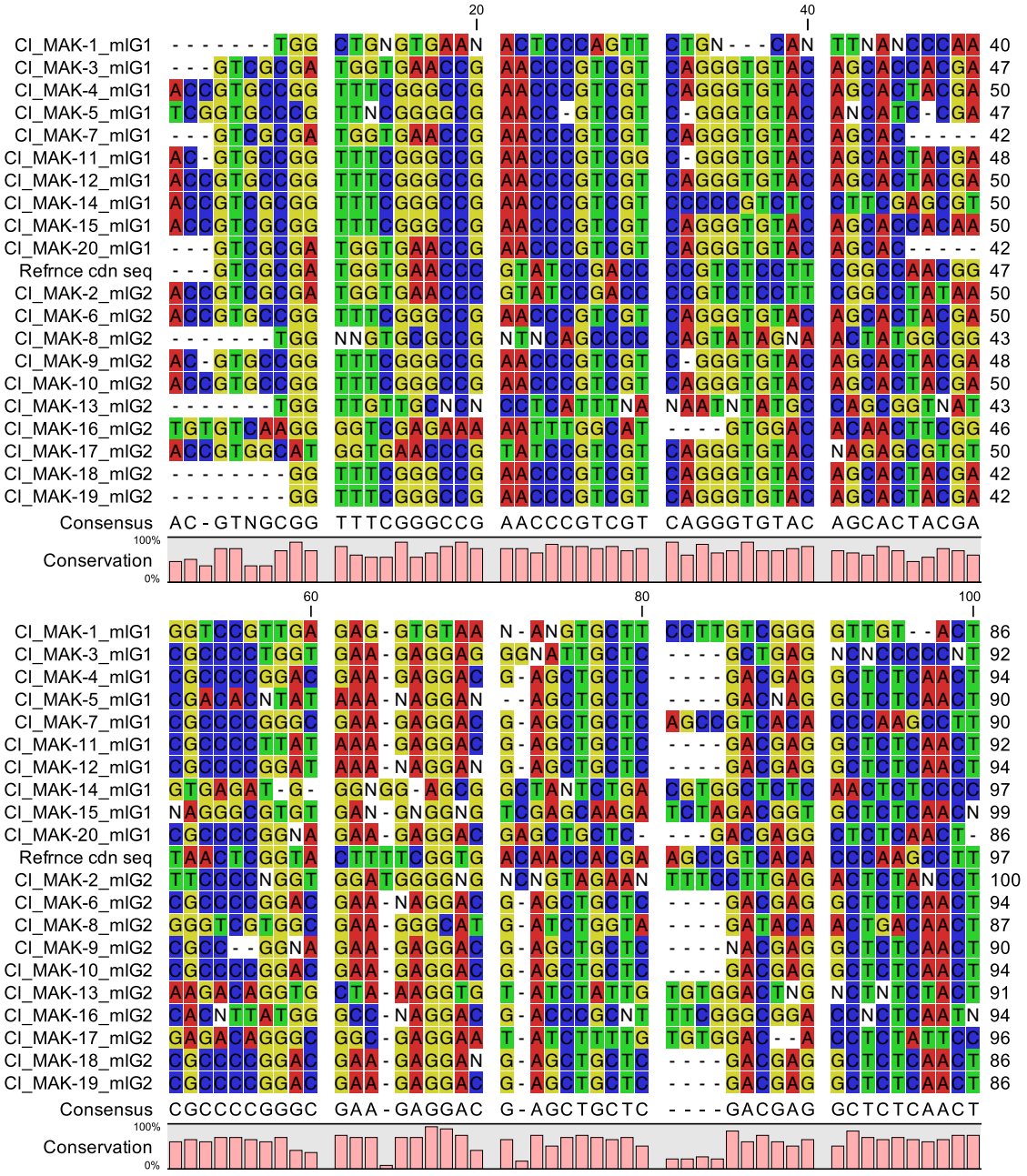
Appendix 6: Sequence alignment of *Cl_HP9* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.

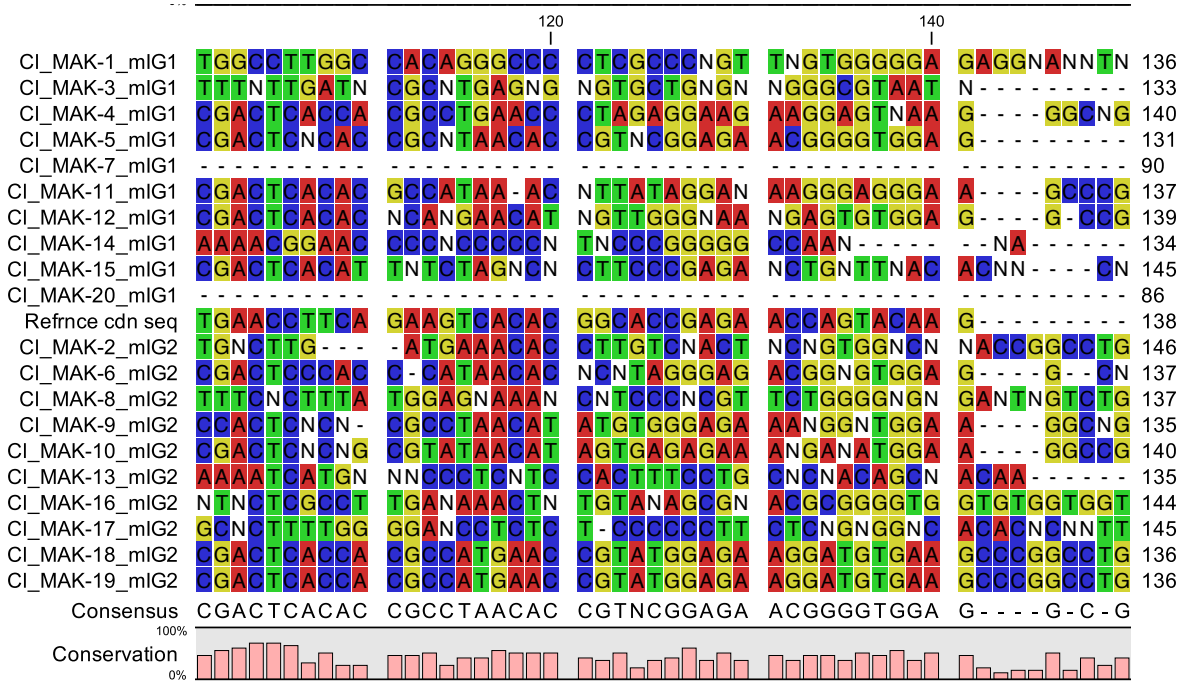
Appendix 7: Sequence alignment of *Cl_IMP* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.



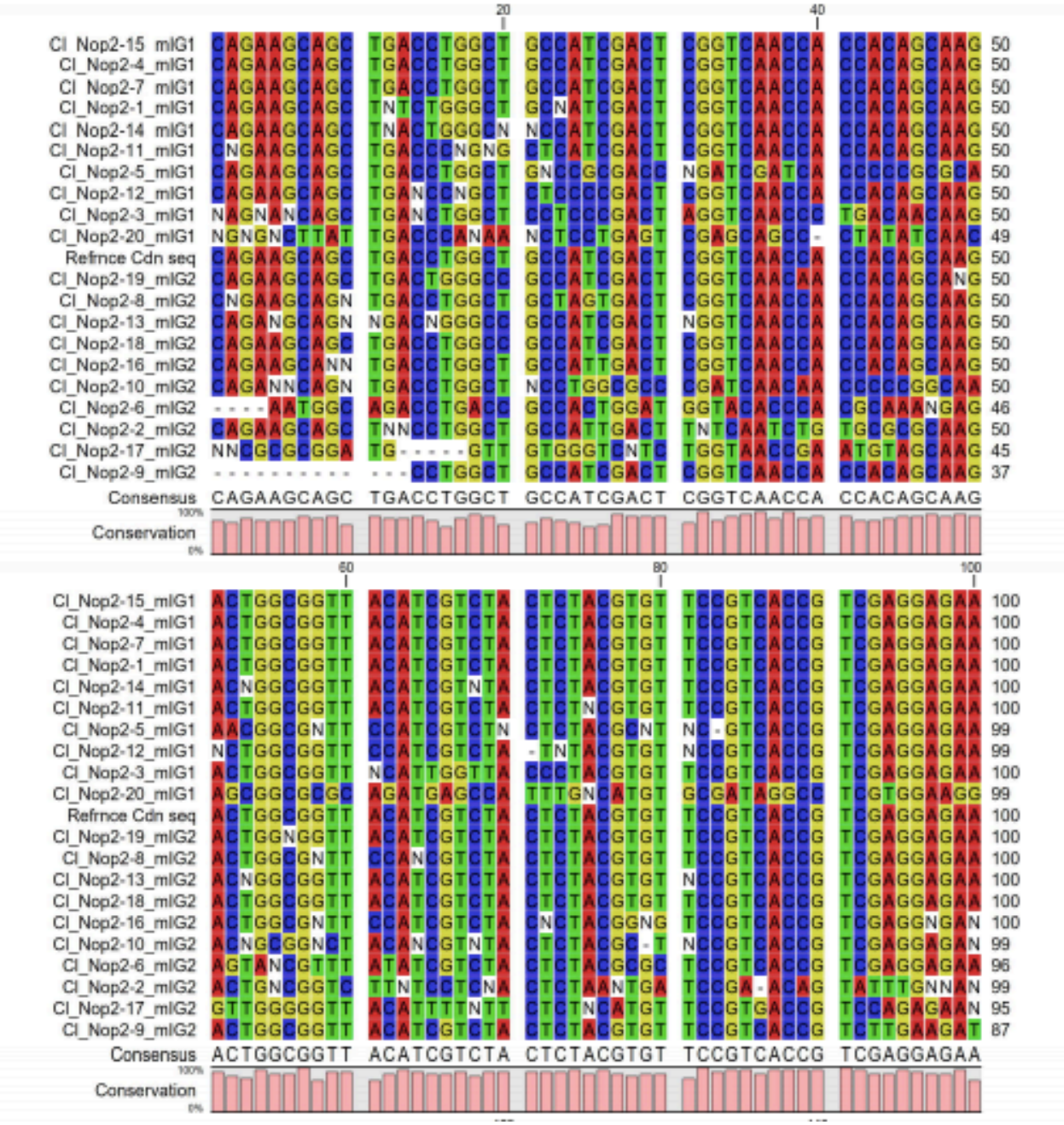

160
180
200
Cl_IMP-1_mlG1 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 160
Cl_IMP-11_mlG1 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 154
Cl_IMP-12_mlG1 CTA-ATAATG-TGTGTTAAA AACACCTGGT GGCGAGTAAG GTGATNNNAC 158
Cl_IMP-14-mlG1 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 156
Cl_IMP-15-mlG1 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 154
Cl_IMP-20_mlG1 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 148
Cl_IMP-3_mlG1 CAA-AGCATG-NGNGAGAAG TACAAGCGGC CNATAGAANG GNNNNCN--- 161
Cl_IMP-4_mlG1 CTA-ANCATG-TGTTAGAAG TACAAANGNC CAATAGAATG G--------- 165
Cl_IMP-5_mlG1 -AAGATGNTG ATGCAANGGN TNNGAGTGGN TGCANCANTN ---------- 162
Cl_IMP-7_mlG1 CAAAAGCNTG ATGACAGAAT GACANTAGGC CTTTCTGGCG TCATATNCTA 166
Cl_IMP-13_mlG2 NGTTGGCACG-T-TGAANAT CCNANTTTGG CGGTANAANC AAACTTTGCT 165
Cl_IMP-10_mlG2 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 148
Cl_IMP-16_mlG2 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG N--------- 154
Cl_IMP-17_mlG2 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 162
Cl_IMP-18_mlG2 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 155
Cl_IMP-19_mlG2 CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 155
Cl_IMP-2_mlG2 CCN-AGCACG-CGTGANAAG TACAGCTGGN CNATNGTNTC GGANTTTGAA 159
Cl_IMP-6_mlG2 CAA-AGCATG-TGNGAGAAG TACAAGTGGC CGATAGAATG GANNNGACAC 159
Cl_IMP-8_mlG2 CAA-AGCATG-TGTGANAAG TACAAGNGGN CNATAGANTC GGNNNNNNNN 159
Cl_IMP-9_mlG2 CAGTTGTACG GCTTGAAAAG C-CAAGNGGG CGATAAAATG GANNGAANCA 168
Refrnce cdn seq CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G--------- 185
Consensus CAA-AGCATG-TGTGAGAAG TACAAGTGGC CGATAGAATG G---------
100%
Conservation
0%

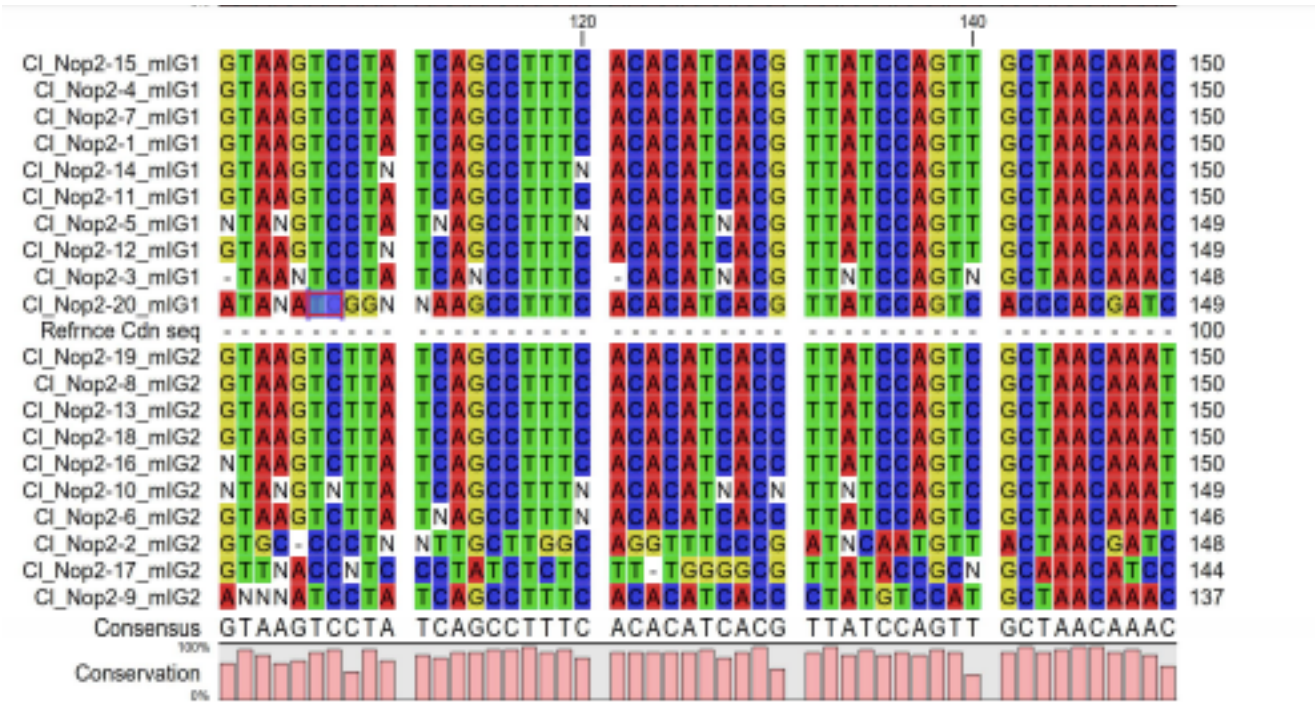
Appendix 8: Sequence alignment of *Cl_MAK* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.

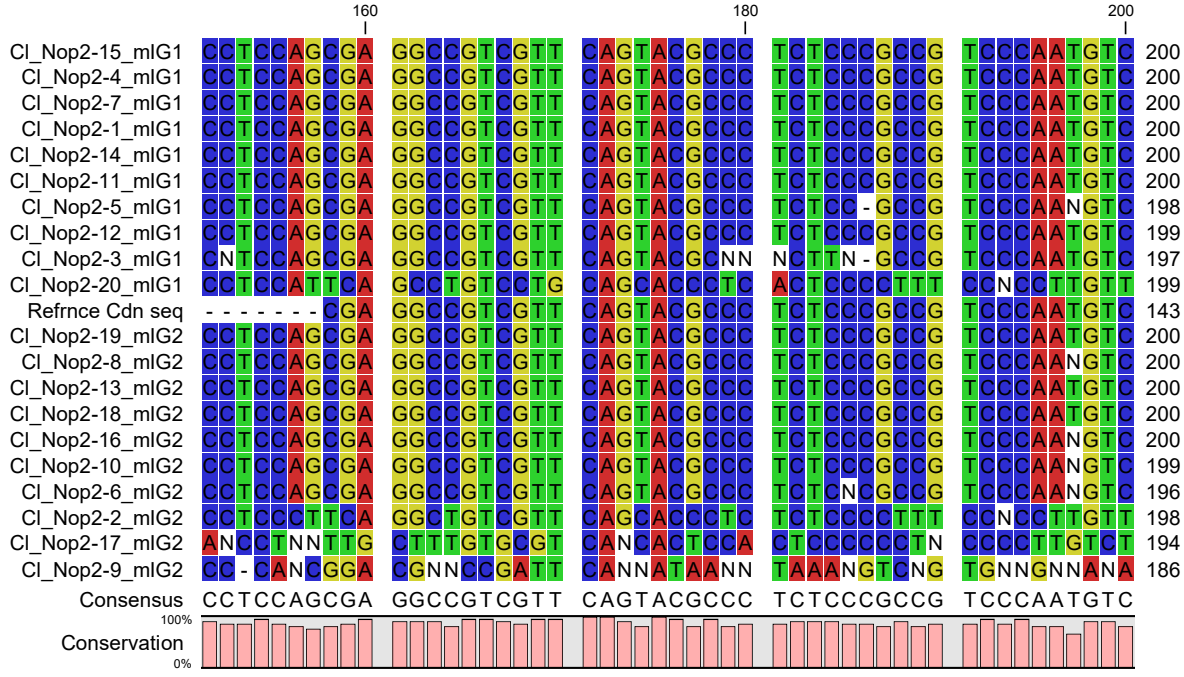



120
140
Cl_MAK-1_mIG1 136
Cl_MAK-3_mIG1 133
Cl_MAK-4_mIG1 140
Cl_MAK-5_mIG1 131
Cl_MAK-7_mIG1 90
Cl_MAK-11_mIG1 137
Cl_MAK-12_mIG1 139
Cl_MAK-14_mIG1 134
Cl_MAK-15_mIG1 145
Cl_MAK-20_mIG1 86
Refrnce cdn seq 138
Cl_MAK-2_mIG2 146
Cl_MAK-6_mIG2 137
Cl_MAK-8_mIG2 137
Cl_MAK-9_mIG2 135
Cl_MAK-10_mIG2 140
Cl_MAK-13_mIG2 135
Cl_MAK-16_mIG2 144
Cl_MAK-17_mIG2 145
Cl_MAK-18_mIG2 136
Cl_MAK-19_mIG2 136
Consensus CGACTCACAC CGCCTAACAC CGTNCGGAGA ACGGGGTGGA G----G-C-G
100%
Conservation
0%

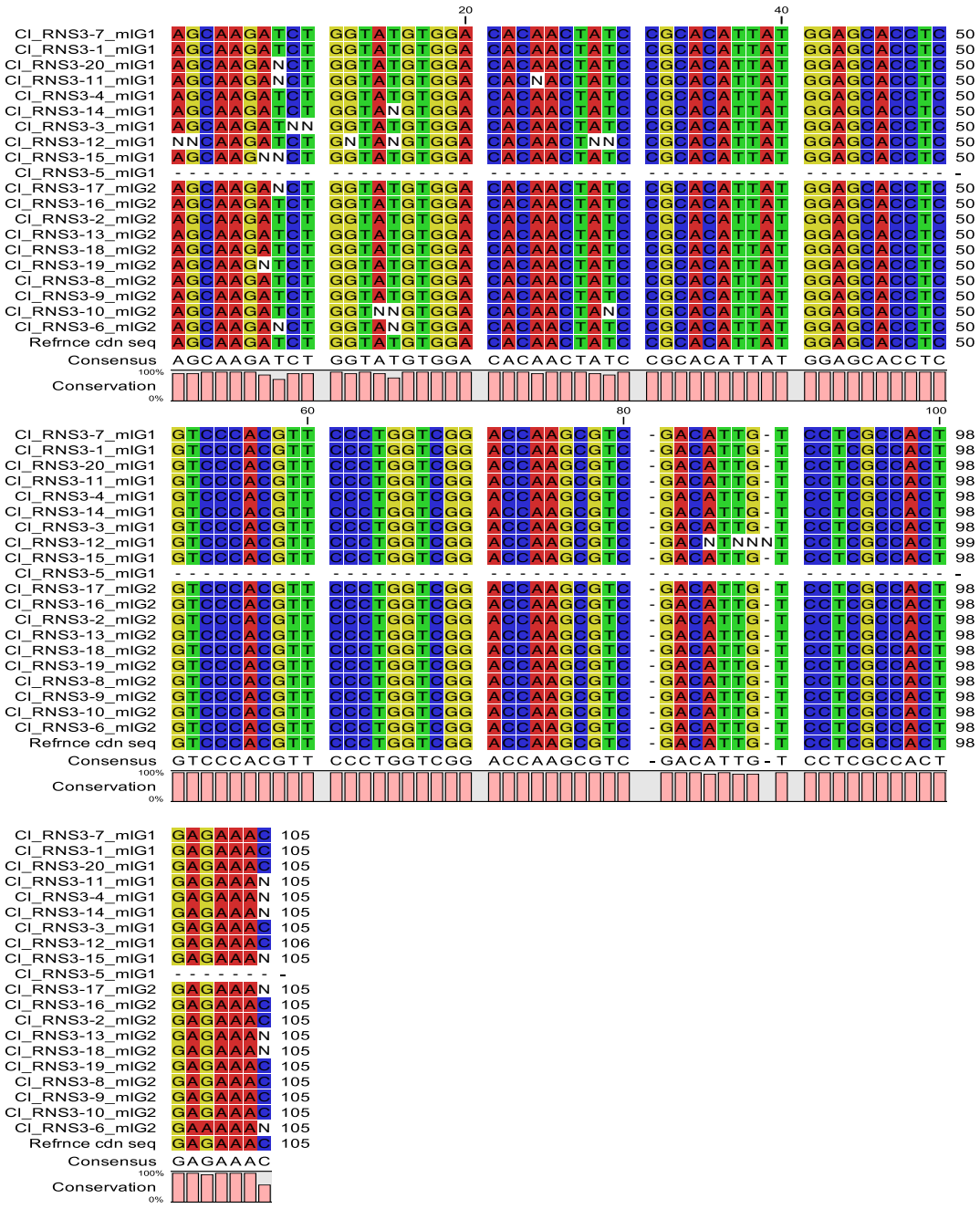
Appendix 9: Sequence alignment of *Cl_Nop2* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.




120
140
Cl_Nop2-15_mlG1 GTAAGTCCTA TCAGCCTTTC ACACATCACG TTATCCAGTT GCTAACAAAC 150
Cl_Nop2-4_mlG1 GTAAGTCCTA TCAGCCTTTC ACACATCACG TTATCCAGTT GCTAACAAAC 150
Cl_Nop2-7_mlG1 GTAAGTCCTA TCAGCCTTTC ACACATCACG TTATCCAGTT GCTAACAAAC 150
Cl_Nop2-1_mlG1 GTAAGTCCTA TCAGCCTTTC ACACATCACG TTATCCAGTT GCTAACAAAC 150
Cl_Nop2-14_mlG1 GTAAGTCCTN TCAGCCTTTN ACACATCACG TTATCCAGTT GCTAACAAAC 150
Cl_Nop2-11_mlG1 GTAAGTCCTA TCAGCCTTTC ACACATCACG TTATCCAGTT GCTAACAAAC 150
Cl_Nop2-5_mlG1 NTANGTCCTA TNAGCCTTTN ACACATNACG TTATCCAGTT GCTAACAAAC 149
Cl_Nop2-12_mlG1 GTAAGTCCTN TCAGCCTTTC ACACATCACG TTATCCAGTT GCTAACAAAC 149
Cl_Nop2-3_mlG1 -TAANTCCTA TCANCCTTTC -CACATNACG TTNTCCAGTN GCTAACAAAC 148
Cl_Nop2-20_mlG1 ATANA GGN NAAGCCTTTC ACACATCACG TTATCCAGTC ACCCACGATC 149
Refrnce Cdn seq 100
Cl_Nop2-19_mlG2 GTAAGTCTTA TCAGCCTTTC ACACATCACC TTATCCAGTC GCTAACAAAT 150
Cl_Nop2-8_mlG2 GTAAGTCTTA TCAGCCTTTC ACACATCACC TTATCCAGTC GCTAACAAAT 150
Cl_Nop2-13_mlG2 GTAAGTCTTA TCAGCCTTTC ACACATCACC TTATCCAGTC GCTAACAAAT 150
Cl_Nop2-18_mlG2 GTAAGTCTTA TCAGCCTTTC ACACATCACC TTATCCAGTC GCTAACAAAT 150
Cl_Nop2-16_mlG2 NTAAGTCTTA TCAGCCTTTC ACACATCACC TTATCCAGTC GCTAACAAAT 150
Cl_Nop2-10_mlG2 NTANGTNTTA TCAGCCTTTN ACACATNACN TTNTCCAGTC GCTAACAAAT 149
Cl_Nop2-6_mlG2 GTAAGTCTTA TNAGCCTTTN ACACATCACC TTATCCAGTC GCTAACAAAT 146
Cl_Nop2-2_mlG2 GTGC-CCCTN NTTGCTTGGC AGGTTTCCCG ATNCAATGTT ACTAACGATC 148
Cl_Nop2-17_mlG2 GTTNACCNTC CGTATCTCTC TT-TGGGGCG TTATACCGCN GCAAACATCC 144
Cl_Nop2-9_mlG2 ANNNATCCTA TCAGCCTTTC ACACATCACC CTATGTCCAT GCTAACAAAC 137
Consensus GTAAGTCCTA TCAGCCTTTC ACACATCACG TTATCCAGTT GCTAACAAAC
100%
Conservation
0%


160
180
200
Cl_Nop2-15_mlG1 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-4_mlG1 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-7_mlG1 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-1_mlG1 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-14_mlG1 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-11_mlG1 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-5_mlG1 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCC-GCCG TCCCAANGTC 198
Cl_Nop2-12_mlG1 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 199
Cl_Nop2-3_mlG1 CNTCCAGCGA GGCCGTCGTT CAGTACGCNN NCTTN-GCCG TCCCAATGTC 197
Cl_Nop2-20_mlG1 CCTCCATTCA GCCTGTCCTG CAGCACCCTC ACTCCCCTTT CCNCCTTGTT 199
Refrnce Cdn seq -------CGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 143
Cl_Nop2-19_mlG2 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-8_mlG2 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAANGTC 200
Cl_Nop2-13_mlG2 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-18_mlG2 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC 200
Cl_Nop2-16_mlG2 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAANGTC 200
Cl_Nop2-10_mlG2 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAANGTC 199
Cl_Nop2-6_mlG2 CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCNCGCCG TCCCAANGTC 196
Cl_Nop2-2_mlG2 CCTCCCTTCA GGCTGTCGTT CAGCACCCTC TCTCCCCCTT CCNCCTTGTT 198
Cl_Nop2-17_mlG2 ANCCTNNTTG CTTTGTGCGT CANCACTCCA CTCCCCCCTN CCCCTTGTCT 194
Cl_Nop2-9_mlG2 CC-CANCGGA CGNNCCGATT CANNATAANN TAAANGTCNG TGNNGNNANA 186
Consensus CCTCCAGCGA GGCCGTCGTT CAGTACGCCC TCTCCCGCCG TCCCAATGTC
100%
Conservation
0%

Appendix 10: Sequence alignment of *Cl_RNS3* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.



Appendix 11: Sequence alignment of *Cl_SART1* that was present in all 20 isolates of two mating incompatibility groups of *Colletotrichum lentis*. These sequences were aligned using multiple sequence alignment tool in CLC genomics software.